

```

; TYPE: PROT
; ORGANISM: Homo sapiens
US-10-322-281-68

Query Match      100.0%; Score 75; DB 16; Length 371;
Best Local Similarity 100.0%; Pred. No. 0.00018;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      1 KILTERGYSPFTTAAE 15
        |||||
Db      191 KILTERGYSPFTTAAE 205

```

Qy	1 KILTERGYSFTTTAE 15 
Dd	191 KILTERGYSFTTTAE 205 

```

RESULT 14
US-10-205-194-93
; Sequence 93, Application US/10205194
; Publication No. US20030134301A1
; GENERAL INFORMATION:
; APPLICANT: Warner-Lambert Company
; APPLICANT: Lee, Kevin
; APPLICANT: Dixon, Alistair
; APPLICANT: Brooksbank, Robert
; APPLICANT: Pinnock, Robert
; TITLE OF INVENTION: Identification and Use of Molecules Implic
; FILE REFERENCE: WL-A-018201
; CURRENT APPLICATION NUMBER: US/10/205,194
; CURRENT FILING DATE: 5200-07-24
; PRIOR APPLICATION NUMBER: GB 0118354.0
; PRIOR FILING DATE: 2001-07-27
; NUMBER OF SEQ ID NOS: 177
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 93
; LENGTH: 375
; TYPE: PRT
; ORGANISM: Rattus rattus
; FEATURE:
; OTHER INFORMATION: Cytoplasmic gamma isoform of actin
US-10-205-194-93

```

```

Query Match      100.0%; Score 75; DB 14; Length 375;
Best Local Similarity 100.0%; Pred. No. 0.00018;
Matches 15; Conservative 0; Mismatches 0; Indels

QY      1 KILTERGYSTTAAE 15
        |||||
Db      191 KILTERGYSTTAAE 205

```

```

RESULT 15
US-10-316-253-88
; Sequence 88, Application US/10316253
; Publication No. US20030162706A1
; GENERAL INFORMATION:
; APPLICANT: The Procter & Gamble Company
; APPLICANT: Peters, Kevin
; APPLICANT: Thompson, Larry
; APPLICANT: Wang, Feng
; APPLICANT: Greis, Kenneth
; TITLE OF INVENTION: Angiogenesis Modulating Proteins
; FILE REFERENCE: 8865M
; CURRENT APPLICATION NUMBER: US/10/316,253
; CURRENT FILING DATE: 2002-12-10
; PRIOR APPLICATION NUMBER: US 60/355,295
; PRIOR FILING DATE: 2002-02-08
; NUMBER OF SEQ ID NOS: 308
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 88
; LENGTH: 375
; TYPE: PRT
; ORGANISM: Rattus norvegicus
US-10-316-253-88

```

Query Match 100.0%; Score 75; DB 14; Length 375;

**THIS PAGE BLANK (USPTO)**

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: April 8, 2005, 09:58:56 ; Search time 9.33333 Seconds  
(without alignments)  
154.634 Million cell updates/sec

Title: US-09-423-351C-9

Perfect score: 76

Sequence: 1 DEAQSKRGILTLKYP 15

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 283416 seqs, 96216763 residues

Total number of hits satisfying chosen parameters: 283416

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

PIR\_79:\*

1: pir1.\*

2: pir2.\*

3: pir3.\*

4: pir4.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	76	100.0	86	2 C43616	actin beta, cytosolic
2	76	100.0	137	2 A28258	actin 5C - fruit fly
3	76	100.0	158	2 I49465	alpha-cardiac actin
4	76	100.0	195	2 S20097	actin 85c - potato
5	76	100.0	213	2 A61043	actin CAL5 - sea s
6	76	100.0	308	2 A03000	actin 3 - fruit fl
7	76	100.0	325	2 J20008	actin homolog prot
8	76	100.0	327	2 S11452	actin (clone 302)
9	76	100.0	328	2 S05430	actin beta - grass
10	76	100.0	349	2 B25819	actin, fetal skele
11	76	100.0	361	2 S68089	actin 2 - Arabidop
12	76	100.0	362	2 A26559	actin type 5, cyto
13	76	100.0	362	2 S68090	actin 8 - Arabidop
14	76	100.0	365	2 A27431	actin, type 1 - Em
15	76	100.0	370	2 A29664	actin - sea urchin
16	76	100.0	374	1 ATBOB	actin beta - bovin
17	76	100.0	374	1 ATBOG	actin gamma - bovi
18	76	100.0	374	2 JC5818	gamma-actin - huma
19	76	100.0	375	1 ATBOSM	actin, aortic smoo
20	76	100.0	375	1 ATRBB	actin, skeletal mu
21	76	100.0	375	1 ATRTC	actin beta - rat
22	76	100.0	375	1 A48324	actin beta, cytosk
23	76	100.0	375	1 ATAX	actin - Acanthamo
24	76	100.0	375	1 ATCHB	actin beta - chick
25	76	100.0	375	1 ATDO	actin - slime mold
26	76	100.0	375	1 ATHUB	actin beta 1 - hu
27	76	100.0	375	1 ATHUG	actin gamma 1 - hu
28	76	100.0	375	1 ATMSB	actin beta - mouse
29	76	100.0	375	1 ATMSG	actin gamma - mous

30	76	100.0	375	1 ATRBB	actin beta, non-mu
31	76	100.0	375	1 S11222	actin gamma, cytos
32	76	100.0	375	2 S33386	actin, cytosolic (
33	76	100.0	375	2 S47897	actin 1 - Pneumocy
34	76	100.0	375	2 T25272	hypothetical prote
35	76	100.0	375	2 A26836	actin - fission ye
36	76	100.0	375	2 S71125	actin beta-2, cyto
37	76	100.0	375	2 S71124	actin beta-1, cyto
38	76	100.0	375	2 S71126	actin beta, cyto
39	76	100.0	375	2 A55001	actin beta - goose
40	76	100.0	375	2 A54728	actin alpha, cardi
41	76	100.0	375	2 S42103	actin - Puccinia g
42	76	100.0	375	2 S70377	actin - Phaffia rh
43	76	100.0	376	1 ATSY3	actin - soybean
44	76	100.0	376	1 A43552	actin gamma, cytos
45	76	100.0	376	1 ATAXE	actin - Entamoeba

## ALIGNMENTS

### RESULT 1

C43616

actin beta, cytosolic - chicken (fragment)

C:Species: Gallus gallus (Chicken)

C>Date: 11-Dec-1992 #sequence\_revision 11-Dec-1992 #text\_change 09-Jul-2004

C:Accession: C43616

R:Paterson, B.M.; Eldridge, J.D.

Science 224, 1436-1438, 1984

A>Title: alpha-Cardiac actin is the major sarcomeric isoform expressed in embryonic avia

A:Reference number: A43616; MUID:84223949; PMID:6729461

A:Accession: C43616

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-86 <PAT>

A:Cross-references: UNIPROT:Q90736; GB:K02173; NID:g211054; PIDN:AAA98513.1; PID:g211055

C:Superfamily: actin

C:Keywords: cytosol; methylated amino acid

F:73/Modified site: 3'-methylhistidine (His) #status predicted

Query Match 100.0%; Score 76; DB 2; Length 86;

Best Local Similarity 100.0%; Pred. No. 7.1e-07;

Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 1 DEAQSKRGILTLKYP 15

Db 56 DEAQSKRGILTLKYP 70

### RESULT 2

A28258

actin 5C - fruit fly (Drosophila melanogaster)

C:Species: Drosophila melanogaster

C>Date: 30-Sep-1989 #sequence\_revision 30-Sep-1989 #text\_change 09-Jul-2004

C:Accession: A28258

R:Vigoreaux, J.O.; Tobin, S.L.

Genes Dev. 1, 1161-1171, 1987

A>Title: Stage-specific selection of alternative transcriptional initiation sites from t

A:Reference number: A28258; MUID:88112795; PMID:3123314

A:Accession: A28258

A:Molecule type: mRNA

A:Residues: 1-137 <VIG>

A:Cross-references: UNIPROT:P10987

A>Note: the authors translated the codon GAG for residue 96 as Gly

C:Genetics:

A:Gene: FlyBase:Act5C

A:Cross-references: FlyBase:FBgn0000042

C:Superfamily: actin

C:Keywords: methylated amino acid

F:74/Modified site: 3'-methylhistidine (His) #status predicted

Query Match 100.0%; Score 76; DB 2; Length 137;

Best Local Similarity 100.0%; Pred. No. 1.1e-06;

Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 DEAQSKRGILTLYKYP 15  
| | | | | | | | | | | | | | | |  
Db 57 DEAQSKRGILTLYKYP 71

RESULT 3  
149465  
alpha-cardiac actin - mouse (fragment)  
C:Species: Mus musculus (house mouse)  
C:Date: 02-Jul-1996 #sequence\_revision 02-Jul-1996 #text\_change 09-Jul-2004  
C:Accession: I49465; I49466  
R:Garner, I.; Minic, A.J.; Alonso, S.; Barton, P.J.R.; Buckingham, M.E.  
EMBO J. 5, 2559-2567, 1986  
A:Title: A 5' duplication of the alpha-cardiac actin gene in BALB/c mice is associated w  
A:Reference number: I49465; MUID:87053822; PMID:3023046  
A:Accession: I49465  
A:Status: translated from GB/EMBL/DBJ  
A:Molecule type: DNA  
A:Residues: 1-158 <RES>  
A:Cross-references: UNIPROT:Q61273; GB:M26775; NID:g191646; PIDN:AAA37165.1; PID:g553858  
A:Accession: I49466  
A:Status: translated from GB/EMBL/DBJ  
A:Molecule type: DNA  
A:Residues: 8-158 <RE2>  
A:Cross-references: GB:M26776; NID:g191649; PIDN:AAA37166.1; PID:g553859  
A:Experimental source: adult cardiac muscle, BALB/c mice  
C:Genetics:  
A:Introns: 50/3  
C:Superfamily: actin  
C:Keywords: cardiac muscle; heart

Query Match 100.0%; Score 76; DB 2; Length 158;  
Best Local Similarity 100.0%; Pred. No. 1.3e-06;  
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 DEAQSKRGILTLYKYP 15  
| | | | | | | | | | | | | | | |  
Db 65 DEAQSKRGILTLYKYP 79

RESULT 4  
S20097  
actin 85c - potato (fragment)  
C:Species: Solanum tuberosum (potato)  
C:Date: 22-Nov-1993 #sequence\_revision 17-Oct-1997 #text\_change 09-Jul-2004  
C:Accession: S20097  
R:Drouin, G.; Dover, G.A.  
J. Mol. Evol. 31, 132-150, 1990  
A:Title: Independent gene evolution in the potato actin gene family demonstrated by phy  
A:Reference number: S20092; MUID:91012599; PMID:2120451  
A:Accession: S20097  
A:Molecule type: DNA  
A:Residues: 1-195 <DRO>  
A:Cross-references: UNIPROT:P30170; EMBL:X55747; NID:g21541; PIDN:CAA39277.1; PID:g13455  
C:Genetics:  
A:Introns: 132/1  
C:Superfamily: actin  
C:Keywords: cytoskeleton; structural protein

Query Match 100.0%; Score 76; DB 2; Length 195;  
Best Local Similarity 100.0%; Pred. No. 1.6e-06;  
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 DEAQSKRGILTLYKYP 15  
| | | | | | | | | | | | | | | |  
Db 38 DEAQSKRGILTLYKYP 52

RESULT 5  
A61043  
actin CA15 - sea squirt (Styela clava) (fragments)

C:Species: Styela clava  
C:Date: 31-Dec-1993 #sequence\_revision 31-Dec-1993 #text\_change 09-Jul-2004  
C:Accession: A61043  
R:Beach, R.L.; Jeffery, W.R.  
Dev. Genet. 11, 2-14, 1990  
A:Title: Temporal and spatial expression of a cytoskeletal actin gene in the ascidian St  
A:Reference number: A61043; MUID:90298580; PMID:2361333  
A:Accession: A61043  
A:Status: not compared with conceptual translation  
A:Molecule type: mRNA  
A:Residues: 1-213 <BEA>  
A:Cross-references: UNIPROT:Q7M3Y7  
C:Comment: This sequence is expressed in cells undergoing rapid cell division.  
C:Superfamily: actin  
C:Keywords: cytoskeleton; methylated amino acid; mitosis; structural protein  
F:73/Modified site: 3'-methylhistidine (His) #status predicted

Query Match 100.0%; Score 76; DB 2; Length 213;  
Best Local Similarity 100.0%; Pred. No. 1.8e-06;  
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 DEAQSKRGILTLYKYP 15  
| | | | | | | | | | | | | | | |  
Db 56 DEAQSKRGILTLYKYP 70

RESULT 6  
A03000  
actin 3 - fruit fly (Drosophila melanogaster) (fragments)  
C:Species: Drosophila melanogaster  
C:Date: 31-Mar-1991 #sequence\_revision 31-Mar-1991 #text\_change 09-Jul-2004  
C:Accession: A03000  
R:Fyrberg, E.A.; Bond, B.J.; Hershey, N.D.; Mixter, K.S.; Davidson, N.  
Cell 24, 107-116, 1981  
A:Title: The actin genes of Drosophila: protein coding regions are highly conserved but  
A:Reference number: A03000; MUID:81210174; PMID:6263481  
A:Accession: A03000  
A:Molecule type: DNA  
A:Residues: 1-308 <FVR>  
A:Cross-references: UNIPROT:P02572  
A:Note: there are 68 unsequenced residues between positions 160 and 161. Partial sequenc  
A:Note: the authors translated the codon GTT for residue 263 as Ile  
C:Genetics:  
A:Gene: FlyBase:Act42A  
A:Cross-references: FlyBase:FBgn0000043  
A:Map position: 42A  
C:Superfamily: actin  
C:Keywords: methylated amino acid  
F:74/Modified site: 3'-methylhistidine (His) #status predicted

Query Match 100.0%; Score 76; DB 2; Length 308;  
Best Local Similarity 100.0%; Pred. No. 2.6e-06;  
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 DEAQSKRGILTLYKYP 15  
| | | | | | | | | | | | | | | |  
Db 57 DEAQSKRGILTLYKYP 71

RESULT 7  
JC2008  
actin homolog protein - red swamp crayfish  
C:Species: Procambarus clarkii (red swamp crayfish)  
C:Date: 14-Jul-1994 #sequence\_revision 14-Jul-1994 #text\_change 20-Jun-2000  
C:Accession: JC2008  
R:Kang, W.K.; Naya, Y.  
Gene 133, 303-304, 1993  
A:Title: Sequence of the cDNA encoding an actin homolog in the crayfish Procambarus cl  
A:Reference number: JC2008; MUID:94040829; PMID:8224920  
A:Accession: JC2008  
A:Molecule type: mRNA  
A:Residues: 1-325 <KAN>  
A:Cross-references: GB:D14612; NID:g434784; PIDN:BAA03463.1; PID:g434785



A;Experimental source: muscle  
C;Superfamily: actin  
C;Keywords: muscle contraction

Query Match 100.0%; Score 76; DB 2; Length 325;  
Best Local Similarity 100.0%; Pred. No. 2.8e-06;  
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 DEAQSKRGILTLYP 15  
|||||  
Db 6 DEAQSKRGILTLYP 20

## RESULT 8

S11452  
actin (clone 302) - brine shrimp (fragment)  
C;Species: Artemia sp. (brine shrimp)  
C;Date: 17-Apr-1993 #sequence\_revision 17-Apr-1993 #text\_change 09-Jul-2004  
C;Accession: S11452  
R;Macias, M.T.; Sastre, L.  
Nucleic Acids Res. 18, 5219-5225, 1990  
A;Title: Molecular cloning and expression of four actin isoforms during Artemia development  
A;Reference number: S11450; MUID:90384823; PMID:2402445  
A;Accession: S11452  
A;Molecule type: mRNA  
A;Residues: 1-327 <MAC>  
A;Cross-references: UNIPROT:P18602; EMBL:X52604; NID:G5666; PIDN:CAA36837.1; PID:G829162  
C;Superfamily: actin  
C;Keywords: methylated amino acid; structural protein  
F;25/Modified site: 3'-methylhistidine (His) #status predicted

Query Match 100.0%; Score 76; DB 2; Length 327;  
Best Local Similarity 100.0%; Pred. No. 2.8e-06;  
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 DEAQSKRGILTLYP 15  
|||||  
Db 8 DEAQSKRGILTLYP 22

## RESULT 9

S05430  
actin beta - grass carp  
C;Species: Ctenopharyngodon idella (grass carp)  
C;Date: 07-Jun-1990 #sequence\_revision 07-Jun-1990 #text\_change 09-Jul-2004  
C;Accession: S05430  
R;Liu, Z.; Zhu, Z.; Roberg, K.; Faras, A.J.; Guise, K.S.; Kapuscinski, A.R.; Hackett, P.  
Nucleic Acids Res. 17, 5850, 1989  
A;Title: The beta-actin gene of carp (Ctenopharyngodon idella).  
A;Reference number: S05430; MUID:89345185; PMID:2762162  
A;Accession: S05430  
A;Status: translation not shown  
A;Molecule type: DNA  
A;Residues: 1-328 <LIU>  
A;Cross-references: UNIPROT:P83751; EMBL:M25013  
C;Genetics:  
A;Introns: 41/3; 121/3; 268/1  
C;Superfamily: actin  
C;Keywords: cytoskeleton; methylated amino acid  
F;73/Modified site: 3'-methylhistidine (His) #status predicted

Query Match 100.0%; Score 76; DB 2; Length 328;  
Best Local Similarity 100.0%; Pred. No. 2.8e-06;  
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 DEAQSKRGILTLYP 15  
|||||  
Db 56 DEAQSKRGILTLYP 70

## RESULT 10

B25819  
actin, fetal skeletal/adult cardiac muscle - mouse (fragment)

C;Species: Mus musculus (house mouse)  
C;Date: 23-Aug-1987 #sequence\_revision 23-Aug-1987 #text\_change 09-Jul-2004  
C;Accession: B25819  
R;Alonso, S.; Minty, A.; Bourlet, Y.; Buckingham, M.  
J. Mol. Evol. 23, 11-22, 1986

A;Title: Comparison of three actin-coding sequences in the mouse; evolutionary relations  
A;Reference number: A25819; MUID:86200234; PMID:3084797

A;Accession: B25819  
A;Molecule type: mRNA  
A;Residues: 1-349 <ALO>

A;Cross-references: UNIPROT:Q61275; GB:X03767; GB:J00381; GB:M10652; NID:G49869; PIDN:CA  
C;Superfamily: actin  
C;Keywords: cardiac muscle; heart; methylated amino acid; muscle; skeletal muscle  
F;47/Modified site: 3'-methylhistidine (His) #status predicted

Query Match 100.0%; Score 76; DB 2; Length 349;  
Best Local Similarity 100.0%; Pred. No. 3e-06;  
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 DEAQSKRGILTLYP 15  
|||||  
Db 30 DEAQSKRGILTLYP 44

## RESULT 11

S68089  
actin 2 - Arabidopsis thaliana  
C;Species: Arabidopsis thaliana (mouse-ear cress)  
C;Date: 05-Dec-1996 #sequence\_revision 13-Mar-1997 #text\_change 09-Jul-2004  
C;Accession: S68089  
R;McDowell, J.M.; Huang, S.; McKinney, E.C.; An, Y.Q.; Meagher, R.B.  
Genetics 142, 587-602, 1996  
A;Title: Structure and evolution of the actin gene family in Arabidopsis thaliana.  
A;Reference number: S68089; MUID:96158109; PMID:8852856  
A;Accession: S68089  
A;Status: not compared with conceptual translation  
A;Molecule type: DNA  
A;Residues: 1-361 <MCD>  
A;Cross-references: UNIPROT:Q93ZL9  
A;Note: mRNA sequencing was also done  
C;Genetics:  
A;Gene: ACT2  
C;Superfamily: actin  
C;Keywords: cytoskeleton; structural protein

Query Match 100.0%; Score 76; DB 2; Length 361;  
Best Local Similarity 100.0%; Pred. No. 3.1e-06;  
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 DEAQSKRGILTLYP 15  
|||||  
Db 53 DEAQSKRGILTLYP 67

## RESULT 12

A26559  
actin type 5, cytosolic - chicken  
C;Species: Gallus gallus (chicken)  
C;Date: 09-Sep-1987 #sequence\_revision 09-Sep-1987 #text\_change 05-Dec-1997  
C;Accession: A26559  
R;Bergsma, D.J.; Chang, K.S.; Schwartz, R.J.  
Mol. Cell. Biol. 5, 1151-1162, 1985  
A;Reference number: A26559; MUID:85213487; PMID:4000121  
A;Accession: A26559  
A;Molecule type: DNA  
A;Residues: 1-362 <BER>  
C;Superfamily: actin  
C;Keywords: cytosol; methylated amino acid  
F;74/Modified site: 3'-methylhistidine (His) #status predicted

Query Match 100.0%; Score 76; DB 2; Length 362;  
Best Local Similarity 100.0%; Pred. No. 3.1e-06;  
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 DEAQSKRGILTLYKYP 15  
 ||||||||||||||||  
 Db 57 DEAQSKRGILTLYKYP 71

## RESULT 13

S68090  
 actin 8 - Arabidopsis thaliana  
 C:Species: Arabidopsis thaliana (mouse-ear cress)  
 C>Date: 05-Dec-1996 #sequence\_revision 13-Mar-1997 #text\_change 09-Jul-2004  
 C:Accession: S68090  
 R:McDowell, J.M.; Huang, S.; McKinney, E.C.; An, Y.Q.; Meagher, R.B.  
 Genetics 142, 597-602, 1996  
 A:Title: Structure and evolution of the actin gene family in Arabidopsis thaliana.  
 A:Reference number: S68089; PMID:96158109; PMID:8852856  
 A:Accession: S68090  
 A>Status: not compared with conceptual translation  
 A:Molecule type: DNA  
 A:Residues: 1-362 <MCD>  
 A:Cross-references: UNIPROT:Q8LB94  
 A:Note: mRNA sequencing was also done  
 C:Genetics:  
 A:Gene: ACT8  
 C:Superfamily: actin  
 C:Keywords: cytoskeleton; structural protein

Query Match 100.0%; Score 76; DB 2; Length 362;  
 Best Local Similarity 100.0%; Pred. No. 3.1e-06;  
 Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 DEAQSKRGILTLYKYP 15  
 ||||||||||||||||  
 Db 54 DEAQSKRGILTLYKYP 68

## RESULT 14

A37431  
 actin, type 1 - Emilia huxleyi (fragment)  
 C:Species: Emilia huxleyi  
 C>Date: 16-Feb-1994 #sequence\_revision 25-Apr-1997 #text\_change 09-Jul-2004  
 C:Accession: A37431  
 R:Bhattacharya, D.; Stickel, S.K.; Sogin, M.L.  
 Mol. Biol. Evol. 10, 689-703, 1993  
 A:Title: Isolation and molecular phylogenetic analysis of actin-coding regions from Emil  
 A:Reference number: A37431; PMID:93330051; PMID:7687735  
 A:Contents: CCMF379  
 A:Accession: A37431  
 A>Status: preliminary  
 A:Molecule type: nucleic acid  
 A:Residues: 1-365 <BHA>  
 A:Cross-references: UNIPROT:Q41205; GB:S64188; NID:9404406; PIDN:AAB27626.1; PID:G404407  
 A:Note: sequence extracted from NCBI backbone (NCBI:135711, NCBI:P:135712)  
 C:Superfamily: actin

Query Match 100.0%; Score 76; DB 2; Length 365;  
 Best Local Similarity 100.0%; Pred. No. 3.2e-06;  
 Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 DEAQSKRGILTLYKYP 15  
 ||||||||||||||||  
 Db 46 DEAQSKRGILTLYKYP 60

## RESULT 15

A29664  
 actin - sea urchin (Strongylocentrotus purpuratus)  
 C:Species: Strongylocentrotus purpuratus (purple urchin)  
 C>Date: 15-Dec-1988 #sequence\_revision 30-Sep-1991 #text\_change 05-Dec-1997  
 C:Accession: A29664  
 R:Crain Jr., W.R.; Boshar, M.F.; Cooper, A.D.; Durica, D.S.; Nagy, A.; Steffen, D.  
 J. Mol. Evol. 25, 37-45, 1987  
 A:Title: The sequence of a sea urchin muscle actin gene suggests a gene conversion with

A:Reference number: A29664; MUID:873111761; PMID:3114500  
 A:Accession: A29664  
 A:Molecule type: DNA  
 A:Residues: 1-370 <CRA>  
 A:Cross-references: GB:X05739; GB:X05740; GB:X05741; GB:X05742; GB:X05743  
 A:Note: the authors translated the codon CAG for residue 260 as Glu  
 C:Superfamily: actin  
 C:Keywords: methylated amino acid  
 F:73/Modified site: 3'-methylhistidine (His) #status predicted  
 Query Match 100.0%; Score 76; DB 2; Length 370;  
 Best Local Similarity 100.0%; Pred. No. 3.2e-06;  
 Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 DEAQSKRGILTLYKYP 15  
 ||||||||||||||||  
 Db 56 DEAQSKRGILTLYKYP 70  
 Search completed: April 8, 2005, 10:53:06  
 Job time : 9.3333 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: April 8, 2005, 09:58:56 ; Search time 57.8667 Seconds  
(without alignments)  
132.739 Million cell updates/sec

Title: US-09-423-351C-9

Perfect score: 76

Sequence: 1 DEAQSKRGILTLPKYP 15

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1612378 seqs, 512079187 residues

Total number of hits satisfying chosen parameters: 1612378

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

UniProt\_03.\*

1: uniprot\_sprot.\*

2: uniprot\_trembl.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	76	100.0	75	2 Q6KB23	Q6KB23 acipenser s
2	76	100.0	78	2 Q61375	Q61375 schmidtea p
3	76	100.0	78	2 Q61376	Q61376 schmidtea p
4	76	100.0	79	2 Q95L83	Q95L83 oryctolagus
5	76	100.0	80	2 Q9TW4	Q9TW4 bos taurus
6	76	100.0	82	2 Q62788	Q62788 sus scrofa
7	76	100.0	82	2 Q9XSX8	Q9XSX8 bubalus bub
8	76	100.0	86	2 Q90736	Q90736 gallus gall
9	76	100.0	91	2 Q7Z211	Q7Z211 salvelinus
10	76	100.0	94	2 Q8MVN0	Q8MVN0 boltonia vi
11	76	100.0	96	2 Q86CV4	Q86CV4 drosophila
12	76	100.0	96	2 Q86FV7	Q86FV7 drosophila
13	76	100.0	100	2 Q862L0	Q862L0 bos taurus
14	76	100.0	104	2 Q86SB7	Q86SB7 gryllus bim
15	76	100.0	108	2 Q18548	Q18548 lytechinus
16	76	100.0	108	2 Q18550	Q18550 lytechinus
17	76	100.0	117	2 Q8T6P7	Q8T6P7 elysia chlo
18	76	100.0	118	2 Q15593	Q15593 entamoeba h
19	76	100.0	121	2 Q8JG71	Q8JG71 oreochromis
20	76	100.0	121	2 Q8JG71	Q8JG71 oreochromis
21	76	100.0	122	2 Q84IA3	Q84IA3 pseudopleur
22	76	100.0	123	2 Q9SSU4	Q9SSU4 cucumis mel
23	76	100.0	129	2 Q6VBE4	Q6VBE4 locusta mlg
24	76	100.0	130	2 Q93298	Q93298 pseudopleur
25	76	100.0	130	2 Q9DE58	Q9DE58 coryphaenoi
26	76	100.0	139	2 Q3DFK2	Q3DFK2 gillichthys
27	76	100.0	144	2 Q90YX9	Q90YX9 fundulus he
28	76	100.0	147	2 Q7X9B5	Q7X9B5 fragaria an
29	76	100.0	148	2 Q8JIZ5	Q8JIZ5 oncorhynch
30	76	100.0	150	2 Q804Y9	Q804Y9 ictalurus p
31	76	100.0	151	2 Q9AVU2	Q9AVU2 ophioglossu

32 76 100.0 151 2 Q61274 mus musculus  
33 76 100.0 151 2 Q9PSV5 Q9PSV5 oryzias lat  
34 76 100.0 155 2 Q9ZQ70 Q9ZQ70 glycyrrhiza  
35 76 100.0 157 2 Q96443 Q96443 glycine max  
36 76 100.0 158 2 Q61273 Q61273 mus musculus  
37 76 100.0 162 2 Q6UBA0 Q6UBA0 chelonis my  
38 76 100.0 164 1 ACT\_SPOLI Q11212 spodoptera  
39 76 100.0 164 2 Q6PX70 Q6PX70 capsicum an  
40 76 100.0 165 2 Q17477 Q17477 haliothis ru  
41 76 100.0 165 2 Q8QFS7 Q8QFS7 scyliorhinu  
42 76 100.0 171 2 Q8JHS9 Q8JHS9 chelydra se  
43 76 100.0 176 2 Q8PZE5 Q8PZE5 brassica na  
44 76 100.0 181 2 Q8JG41 Q8JG41 dicentrarch  
45 76 100.0 181 2 Q8AWC8 Q8AWC8 chionodraco

#### ALIGNMENTS

RESULT 1  
Q6KB23 PRELIMINARY; PRT; 75 AA.  
AC Q6KB23;  
DT 05-JUL-2004 (Tremblrel. 27, Created)  
DT 05-JUL-2004 (Tremblrel. 27, Last sequence update)  
DT 05-JUL-2004 (Tremblrel. 27, Last annotation update)  
DE Beta-actin (Fragment).  
OS Acipenser sinensis (Chinese sturgeon).  
OG Chloroplast.  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Actinopterygii; Chondrostei; Acipenseriformes; Acipenseridae;  
OC Acipenser.  
OX NCBI\_TaxID=61970;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC TISSUE=Liver;  
RA Zhang Z.;  
RL Submitted (JUN-2004) to the EMBL/GenBank/DBJ databases.  
[2]  
RP SEQUENCE FROM N.A.  
RC TISSUE=Liver;  
RA Zhaobin Z.;  
RL Submitted (JUN-2004) to the EMBL/GenBank/DBJ databases.  
DR EMBL; AJ745100; CAG33738.1; -  
DR GO; GO:0005884; C:actin filament; IEA.  
DR GO; GO:0009507; C:chloroplast; IEA.  
DR GO; GO:0003774; F:motor activity; IEA.  
DR GO; GO:0005200; F:structural constituent of cytoskeleton; IEA.  
DR InterPro; IPR004001; Actin.  
DR PRINTS; PR00190; ACTIN.  
DR PROSITE; PS00406; ACTINS\_1; 1.  
KW Chloroplast.  
FT NON TER 1  
FT NON TER 75  
SQ SEQUENCE 75 AA; 8545 MW; 32F1B8CC7C40E388 CRC64;

Query Match 100.0%; Score 76; DB 2; Length 75;  
Best Local Similarity 100.0%; Pred. No. 2e-06;  
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 DEAQSKRGILTLPKYP 15  
| | | | | | | | | | | | | | | | | |  
Db 16 DEAQSKRGILTLPKYP 30

RESULT 2  
Q61375 PRELIMINARY; PRT; 78 AA.  
ID Q61375  
AC Q61375;  
DT 01-AUG-1998 (Tremblrel. 07, Created)  
DT 01-AUG-1998 (Tremblrel. 07, Last sequence update)  
DT 01-JUN-2003 (Tremblrel. 24, Last annotation update)

DE Actin 1 (Fragment).  
GN Name=DpAct1;  
OS Schmidtea polychroa.  
OC Eukaryota; Metazoa; Platyhelminthes; Turbellaria; Seriata; Tricladida;  
OC Paludicola; Dugesidae; Schmidtea.  
OX NCBI\_TaxID=50054;  
RN (1)  
RP SEQUENCE FROM N.A.  
RX MEDLINE=98186847; PubMed=9518500; DOI=10.1093/nar/26.8.2031;  
RA Fagotti A., Gabbiani G., Pascolini R., Neuville P.;  
RT "Multiple isoform recovery (MIR)-PCR: a simple method for the  
RT isolation of related mRNA isoforms.";  
RL Nucleic Acids Res. 26:2031-2033(1998).  
CC -!- FUNCTION: Actins are highly conserved proteins that are involved  
CC in various types of cell motility and are ubiquitously expressed  
CC in all eukaryotic cells (By similarity).  
CC -!- SIMILARITY: Belongs to the actin family.  
CC EMBL; AF027161; AAC38981.1; -.  
DR HSSP; P02577; INM1.  
DR GO; GO:0005884; C:actin filament; IEA.  
DR GO; GO:0003774; F:motor activity; IEA.  
DR GO; GO:0005200; F:structural constituent of cytoskeleton; IEA.  
DR InterPro; IPR004001; Actin.  
DR Pfam; PF00022; Actin; 1.  
DR PRINTS; PR00190; ACTIN.  
DR PROSITE; PS00406; ACTINS\_1; 1.  
DR Structural protein. 78  
KW NON TER  
FT SEQUENCE 78 AA; 8205 MW; 572BAB2698C24404 CRC64;  
SQ

Query Match 100.0%; Score 76; DB 2; Length 78;  
Best Local Similarity 100.0%; Pred. No. 2e-06;  
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 DEAQSKRGILTLYP 15  
Db 57 DEAQSKRGILTLYP 71

RESULT 3  
O61376 PRELIMINARY; PRT; 78 AA.  
ID O61376;  
AC O61376;  
DT 01-AUG-1998 (TrEMBLrel. 07, Created)  
DT 01-AUG-1998 (TrEMBLrel. 07, Last sequence update)  
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)  
DE Actin 2 (Fragment).  
GN Name=DpAct2;  
OS Schmidtea polychroa.  
OC Eukaryota; Metazoa; Platyhelminthes; Turbellaria; Seriata; Tricladida;  
OC Paludicola; Dugesidae; Schmidtea.  
OX NCBI\_TaxID=50054;  
RN (1)  
RP SEQUENCE FROM N.A.  
RX MEDLINE=98186847; PubMed=9518500; DOI=10.1093/nar/26.8.2031;  
RA Fagotti A., Gabbiani G., Pascolini R., Neuville P.;  
RT "Multiple isoform recovery (MIR)-PCR: a simple method for the  
RT isolation of related mRNA isoforms.";  
RL Nucleic Acids Res. 26:2031-2033(1998).  
CC -!- FUNCTION: Actins are highly conserved proteins that are involved  
CC in various types of cell motility and are ubiquitously expressed  
CC in all eukaryotic cells (By similarity).  
CC -!- SIMILARITY: Belongs to the actin family.  
CC EMBL; AF027162; AAC38982.1; -.  
DR HSSP; P02577; INM1.  
DR GO; GO:0005884; C:actin filament; IEA.  
DR GO; GO:0003774; F:motor activity; IEA.  
DR GO; GO:0005200; F:structural constituent of cytoskeleton; IEA.  
DR InterPro; IPR004001; Actin.  
DR Pfam; PF00022; Actin; 1.  
DR PRINTS; PR00190; ACTIN.

DR PROSITE; PS00406; ACTINS\_1; 1.  
KW Structural protein.  
FT NON TER 78  
SQ SEQUENCE 78 AA; 8265 MW; B9DFF9B28E4ADB0E CRC64;  
Query Match 100.0%; Score 76; DB 2; Length 78;  
Best Local Similarity 100.0%; Pred. No. 2e-06;  
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 DEAQSKRGILTLYP 15  
Db 57 DEAQSKRGILTLYP 71

RESULT 4  
Q95L83 PRELIMINARY; PRT; 79 AA.  
ID Q95L83;  
AC Q95L83;  
DT 01-DEC-2001 (TrEMBLrel. 19, Created)  
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)  
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)  
DE Beta-actin (Fragment).  
OS Oryctolagus cuniculus (Rabbit).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Lagomorpha; Leporidae; Oryctolagus.  
OX NCBI\_TaxID=9986;  
RN (1)  
RP SEQUENCE FROM N.A.  
RX MEDLINE=22306422; PubMed=12388084;  
RA Wang D.-A., Du H., Jaggar J.H., Brindley D.N., Tigyi G.J.,  
RA Watsky M.A.;  
RT "Injury-elicited differential transcriptional regulation of  
RT phospholipid growth factor receptors in the cornea";  
RL Am. J. Physiol. Cell Physiol. 283:C1646-C1654(2002).  
CC -!- FUNCTION: Actins are highly conserved proteins that are involved  
CC in various types of cell motility and are ubiquitously expressed  
CC in all eukaryotic cells (By similarity).  
CC -!- SIMILARITY: Belongs to the actin family.  
CC EMBL; AF404278; AAL01885.1; -.  
DR HSSP; P02577; INM1.  
DR GO; GO:0005884; C:actin filament; IEA.  
DR GO; GO:0003774; F:motor activity; IEA.  
DR GO; GO:0005200; F:structural constituent of cytoskeleton; IEA.  
DR InterPro; IPR004001; Actin.  
DR InterPro; IPR004000; Actin\_like.  
DR Pfam; PF00022; Actin; 1.  
DR PRINTS; PR00190; ACTIN.  
DR PROSITE; PS00406; ACTINS\_1; 1.  
KW Structural protein.  
FT NON TER 79  
SQ SEQUENCE 79 AA; 8420 MW; 05056E4E2AC092F8 CRC64;  
Query Match 100.0%; Score 76; DB 2; Length 79;  
Best Local Similarity 100.0%; Pred. No. 2.1e-06;  
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 DEAQSKRGILTLYP 15  
Db 58 DEAQSKRGILTLYP 72

RESULT 5  
Q9TWT4 PRELIMINARY; PRT; 80 AA.  
ID Q9TWT4;  
AC Q9TWT4;  
DT 01-MAY-2000 (TrEMBLrel. 13, Created)  
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)  
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)  
DE Beta actin (Fragment).  
OS Bos taurus (Bovine).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;

```
OC Bovinae; Bos.
OX NCBI_TaxID=9913;
RN [1]
RP SEQUENCE FROM N.A.
RA John S.J., Bilodeau-Goeseels S.; to the EMBL/GenBank/DBJ databases.
RL Submitted (OCT-1999) to the EMBL/GenBank/DBJ databases.
CC -!- FUNCTION: Actins are highly conserved proteins that are involved
CC in various types of cell motility and are ubiquitously expressed
CC in all eukaryotic cells (By similarity).
CC -!- SIMILARITY: Belongs to the actin family.
DR EMBL; AF191490; AAF05984.1; -.
DR HSSP; P02577; INWI.
DR GO; GO:0005884; C:actin filament; IEA.
DR GO; GO:0003774; F:motor activity; IEA.
DR GO; GO:0005200; F:structural constituent of cytoskeleton; IEA.
DR InterPro; IPR004001; Actin.
DR InterPro; IPR004000; Actin_like.
DR Pfam; PF00022; Actin; 1.
DR PRINTS; PR00190; ACTIN.
DR PROSITE; PS00406; ACTINS_1; 1.
KW Structural protein.
FT NON_TER 1
FT NON_TER 80
SQ SEQUENCE 80 AA; 9133 MW; 47354ABB7010668D CRC64;

Query Match 100.0%; Score 76; DB 2; Length 80;
Best Local Similarity 100.0%; Pred. No. 2.1e-06;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 1 DEAQSKRGILTLYKYP 15
Db 22 DEAQSKRGILTLYKYP 36

RESULT 6
O62788 PRELIMINARY; PRT; 82 AA.
AC O62788;
DT 01-AUG-1998 (TREMBLrel. 07, Created)
DT 01-AUG-1998 (TREMBLrel. 07, Last sequence update)
DT 01-JUN-2003 (TREMBLrel. 24, Last annotation update)
DE Beta actin (Fragment).
OS Sus scrofa (Pig).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.
OX NCBI_TaxID=9823;
RN [1]
RP SEQUENCE FROM N.A.
RA McNeel R.L., Mersmann H.J.;
RL "Distribution and quantification of beta1-, beta2-, and beta3-
RT adrenergic receptor subtype transcripts in porcine tissues.";
RL J. Anim. Sci. 77:611-621(1999).
CC -!- FUNCTION: Actins are highly conserved proteins that are involved
CC in various types of cell motility and are ubiquitously expressed
CC in all eukaryotic cells (By similarity).
CC -!- SIMILARITY: Belongs to the actin family.
DR EMBL; AF054837; AAC08590.1; -.
DR HSSP; P02568; ILCU.
DR GO; GO:0005884; C:actin filament; IEA.
DR GO; GO:0003774; F:motor activity; IEA.
DR GO; GO:0005200; F:structural constituent of cytoskeleton; IEA.
DR InterPro; IPR004001; Actin.
DR InterPro; IPR004000; Actin_like.
DR Pfam; PF00022; Actin; 1.
DR PRINTS; PR00190; ACTIN.
DR PROSITE; PS00406; ACTINS_1; 1.
DR PROSITE; PS01132; ACTINS_ACT_LIKE; 1.
KW Structural protein.
FT NON_TER 1
FT NON_TER 82
SQ SEQUENCE 82 AA; 9410 MW; 8FCC5A5BDAB9F9DC CRC64;
```

```
Query Match 100.0%; Score 76; DB 2; Length 82;
Best Local Similarity 100.0%; Pred. No. 2.1e-06;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 1 DEAQSKRGILTLYKYP 15
Db 22 DEAQSKRGILTLYKYP 36

RESULT 7
O9XSX8 PRELIMINARY; PRT; 82 AA.
AC O9XSX8;
DT 01-NOV-1999 (TREMBLrel. 12, Created)
DT 01-NOV-1999 (TREMBLrel. 12, Last sequence update)
DT 01-JUN-2003 (TREMBLrel. 24, Last annotation update)
DE Beta actin (Fragment)..
OS Bubalus bubalis (Domestic water buffalo).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
OC Bovinae; Bubalus.
OX NCBI_TaxID=89462;
RN [1]
RP SEQUENCE FROM N.A.
RA Chattopadhyay M., Gangadharan S., Ali S.;
RL Submitted (JUN-1999) to the EMBL/GenBank/DBJ databases.
CC -!- FUNCTION: Actins are highly conserved proteins that are involved
CC in various types of cell motility and are ubiquitously expressed
CC in all eukaryotic cells (By similarity).
CC -!- SIMILARITY: Belongs to the actin family.
DR EMBL; AF156978; AAD40181.1; -.
DR HSSP; P02577; INWI.
DR GO; GO:0005884; C:actin filament; IEA.
DR GO; GO:0003774; F:motor activity; IEA.
DR GO; GO:0005200; F:structural constituent of cytoskeleton; IEA.
DR InterPro; IPR004001; Actin.
DR InterPro; IPR004000; Actin_like.
DR Pfam; PF00022; Actin; 1.
DR PRINTS; PR00190; ACTIN.
DR PROSITE; PS00406; ACTINS_1; 1.
DR PROSITE; PS01132; ACTINS_ACT_LIKE; 1.
KW Structural protein.
FT NON_TER 1
FT NON_TER 82
SQ SEQUENCE 82 AA; 9417 MW; 9709C6112FAF296C CRC64;

Query Match 100.0%; Score 76; DB 2; Length 82;
Best Local Similarity 100.0%; Pred. No. 2.1e-06;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 1 DEAQSKRGILTLYKYP 15
Db 22 DEAQSKRGILTLYKYP 36

RESULT 8
Q90736 PRELIMINARY; PRT; 86 AA.
AC Q90736;
DT 01-NOV-1996 (TREMBLrel. 01, Created)
DT 01-NOV-1996 (TREMBLrel. 01, Last sequence update)
DT 01-JUN-2003 (TREMBLrel. 24, Last annotation update)
DE Beta-actin (Fragment)..
OS Gallus gallus (Chicken).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Archosauria; Aves; Neognathae; Galliformes; Phasianinae;
OC Gallus.
OX NCBI_TaxID=9031;
RN [1]
RP SEQUENCE FROM N.A.
RA Paterson B.M., Eldridge J.D.;
RX MEDLINE=84223949; PubMed=6729461;
```

RT "alpha-Cardiac actin is the major sarcomeric isoform expressed in  
 RL embryonic avian skeletal muscle.";  
 CC Science 224:1436-1438(1984).  
 CC -!- FUNCTION: Actins are highly conserved proteins that are involved  
 CC in various types of cell motility and are ubiquitously expressed  
 CC in all eukaryotic cells (By similarity).  
 CC -!- SIMILARITY: Belongs to the actin family.  
 DR EMBL: K02173; AA98513.1; -.  
 DR PIR: C43616; C43616.  
 DR HSP: P02577; INM1.  
 DR GO: GO:0005884; C:actin filament; IEA.  
 DR GO: GO:0003774; F:motor activity; IEA.  
 DR GO: GO:0005200; F:structural constituent of cytoskeleton; IEA.  
 DR Pfam: PF00022; Actin; 1.  
 DR PRINTS: PR00190; ACTIN.  
 DR PROSITE: PS00406; ACTINS\_1; 1.  
 DR Structural protein.  
 FT NON TER 86  
 SQ SEQUENCE 86 AA; 9351 MW; A55285196A328E6E CRC64;

Query Match 100.0%; Score 76; DB 2; Length 86;  
 Best Local Similarity 100.0%; Pred. No. 2.3e-06;  
 Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 DEAQSKRGILTLYP 15  
 |||||  
 Db 56 DEAQSKRGILTLYP 70

## RESULT 9

ID Q72Z11 PRELIMINARY; PRT; 91 AA.  
 AC Q72Z11;  
 DT 01-JUN-2003 (TrEMBLrel. 24, Created)  
 DT 01-JUN-2003 (TrEMBLrel. 24, Last sequence update)  
 DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)  
 DE Beta-actin (Fragment).  
 OS Salvelinus alpinus (Arctic char).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Actinopterygii; Neopterygii; Teleostei; Euteleostei;  
 OC Protacanthopterygii; Salmoniformes; Salmonidae; Salvelinus.  
 OX NCBI\_TaxID=8036;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA McGowan C., Davidson E.A., Davidson W.S.;  
 RL Submitted (MAR-2003) to the EMBL/GenBank/DBJ databases.  
 CC -!- FUNCTION: Actins are highly conserved proteins that are involved  
 CC in various types of cell motility and are ubiquitously expressed  
 CC in all eukaryotic cells (By similarity).  
 CC -!- SIMILARITY: Belongs to the actin family.  
 DR EMBL: AY262761; AAP31127.1; -.  
 DR HSP: P02577; INM1.  
 DR GO: GO:0005884; C:actin filament; IEA.  
 DR GO: GO:0003774; F:motor activity; IEA.  
 DR GO: GO:0005200; F:structural constituent of cytoskeleton; IEA.  
 DR InterPro: IPR004001; Actin.  
 DR InterPro: IPR004000; Actin\_like.  
 DR Pfam: PF00022; Actin; 1.  
 DR PRINTS: PR00190; ACTIN.  
 DR PROSITE: PS00406; ACTINS\_1; 1.  
 DR Structural protein.  
 FT NON TER 91  
 SQ SEQUENCE 91 AA; 10063 MW; 7118DB6663CD895C CRC64;

Query Match 100.0%; Score 76; DB 2; Length 91;  
 Best Local Similarity 100.0%; Pred. No. 2.4e-06;  
 Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 DEAQSKRGILTLYP 15  
 |||||  
 Db 54 DEAQSKRGILTLYP 69

## RESULT 10

ID Q8MVNO PRELIMINARY; PRT; 94 AA.  
 AC Q8MVNO;  
 DT 01-OCT-2002 (TrEMBLrel. 22, Created)  
 DT 01-OCT-2002 (TrEMBLrel. 22, Last sequence update)  
 DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)  
 DE Cytoskeletal actin 2-like protein (Fragment).  
 DR Name=ca2;  
 OS Bolitena villosa.  
 OC Eukaryota; Metazoa; Chordata; Urochordata; Ascidiacea;  
 OC Stolidobranchia; Pyuridae; Bolitena.  
 OX NCBI\_TaxID=63515;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=22248966; PubMed=12361966;  
 RA Davidson B.J., Swalla B.J.;  
 RT "A molecular analysis of ascidian metamorphosis reveals activation of  
 RT an innate immune response";  
 RL Development 129:4739-4751(2002).  
 CC -!- FUNCTION: Actins are highly conserved proteins that are involved  
 CC in various types of cell motility and are ubiquitously expressed  
 CC in all eukaryotic cells (By similarity).  
 CC -!- SIMILARITY: Belongs to the actin family.  
 DR EMBL: AF483031; AAM76111.1; -.  
 DR HSP: P02577; INM1.  
 DR GO: GO:0005884; C:actin filament; IEA.  
 DR GO: GO:0003774; F:motor activity; IEA.  
 DR GO: GO:0005200; F:structural constituent of cytoskeleton; IEA.  
 DR InterPro: IPR004001; Actin.  
 DR InterPro: IPR004000; Actin\_like.  
 DR Pfam: PF00022; Actin; 1.  
 DR PRINTS: PR00190; ACTIN.  
 DR PROSITE: PS00406; ACTINS\_1; 1.  
 DR Structural protein.  
 FT NON TER 94  
 SQ SEQUENCE 94 AA; 10576 MW; 6CEAB58DA12372A4 CRC64;

Query Match 100.0%; Score 76; DB 2; Length 94;  
 Best Local Similarity 100.0%; Pred. No. 2.5e-06;  
 Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 DEAQSKRGILTLYP 15  
 |||||  
 Db 41 DEAQSKRGILTLYP 55

## RESULT 11

ID Q86CV4 PRELIMINARY; PRT; 96 AA.  
 AC Q86CV4;  
 DT 01-JUN-2003 (TrEMBLrel. 24, Created)  
 DT 01-JUN-2003 (TrEMBLrel. 24, Last sequence update)  
 DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)  
 DE Actin E2 (Fragment).  
 CN Name=ActE2;  
 OS Drosophila novaezicana.  
 OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;  
 OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;  
 OC Ephydroidea; Drosophilidae; Drosophila.  
 OX NCBI\_TaxID=47314;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=15010-1031.0;  
 RX MEDLINE=22480340; PubMed=12574518; DOI=10.1073/pnas.0336368100;  
 RA Wittkopp P.J., Williams B.L., Selegue J.E., Carroll S.B.;  
 RT "Drosophila pigmentation evolution: divergent genotypes underlying  
 RT convergent phenotypes";  
 RL Proc. Natl. Acad. Sci. U.S.A. 100:1808-1813(2003).  
 CC -!- FUNCTION: Actins are highly conserved proteins that are involved  
 CC in various types of cell motility and are ubiquitously expressed  
 CC in all eukaryotic cells (By similarity).

-!- SIMILARITY: Belongs to the actin family.  
 EMBL; AY165541; AAP21565.1; -.  
 HSSP; P02568; 1LCU.  
 FlyBase; FBgn0066214; Dnov\ActE2.  
 GO; GO:0005884; C:actin filament; IEA.  
 GO; GO:0003774; F:motor activity; IEA.  
 GO; GO:0005200; F:structural constituent of cytoskeleton; IEA.  
 InterPro; IPR004001; Actin.  
 InterPro; IPR004001; Actin\_like.  
 Pfam; PF00022; Actin; 1.  
 PRINTS; PR00190; ACTIN.  
 PROSITE; PS00406; ACTINS\_1; 1.  
 Structural protein.  
 NON\_TER 96  
 SEQUENCE 96 AA; 10610 MW; 1454B7D9DD3F3564 CRC64;

Query Match 100.0%; Score 76; DB 2; Length 96;  
 Best Local Similarity 100.0%; Pred. No. 2.5e-06;  
 Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 DEAQSKRGILTLYP 15  
 |||||  
 Db 57 DEAQSKRGILTLYP 71

RESULT 12  
 Q86FV7 PRELIMINARY; PRT; 96 AA.  
 ID Q86FV7  
 AC Q86FV7  
 DT 01-JUN-2003 (TrEMBLrel. 24, Created)  
 DT 01-JUN-2003 (TrEMBLrel. 24, Last sequence update)  
 DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)  
 Actin E2 (Fragment).  
 Name=ActE2;  
 GN Drosophila americana (Fruit fly).  
 OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;  
 OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;  
 OC Ephydroidea; Drosophilidae; Drosophila.  
 OX NCBI\_TaxID=40366;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=15010-0951.0;  
 RX MEDLINE=22480340; PubMed=12574518; DOI=10.1073/pnas.0336368100;  
 RA Wittkopp P.J., Williams B.L., Selegue J.E., Carroll S.B.;  
 RT "Drosophila pigmentation evolution: divergent genotypes underlying  
 convergent phenotypes".  
 RL Proc. Natl. Acad. Sci. U.S.A. 100:1808-1813(2003).  
 CC -!- FUNCTION: Actins are highly conserved proteins that are involved  
 in various types of cell motility and are ubiquitously expressed  
 in all eukaryotic cells (By similarity).  
 CC -!- SIMILARITY: Belongs to the actin family.

EMBL; AY165516; AAP12711.1; -.  
 HSSP; P02568; 1LCU.  
 FlyBase; FBgn0066281; Dame\ActE2.  
 GO; GO:0005884; C:actin filament; IEA.  
 GO; GO:0003774; F:motor activity; IEA.  
 GO; GO:0005200; F:structural constituent of cytoskeleton; IEA.  
 InterPro; IPR004001; Actin.  
 InterPro; IPR004000; Actin\_like.  
 Pfam; PF00022; Actin; 1.  
 PRINTS; PR00190; ACTIN.  
 PROSITE; PS00406; ACTINS\_1; 1.  
 Structural protein.  
 NON\_TER 96  
 SEQUENCE 96 AA; 10610 MW; 1454B7D9DD3F3564 CRC64;

Query Match 100.0%; Score 76; DB 2; Length 96;  
 Best Local Similarity 100.0%; Pred. No. 2.5e-06;  
 Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 DEAQSKRGILTLYP 15  
 |||||  
 Db 57 DEAQSKRGILTLYP 71

RESULT 13  
 Q862L0 PRELIMINARY; PRT; 100 AA.  
 ID Q862L0  
 AC Q862L0  
 DT 01-JUN-2003 (TrEMBLrel. 24, Created)  
 DT 01-JUN-2003 (TrEMBLrel. 24, Last sequence update)  
 DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)  
 Similar to b-actin (Fragment).  
 DE Bos taurus (Bovine).  
 OS Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;  
 OC Bovinae; Bos.  
 OX NCBI\_TaxID=9913;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=22544902; PubMed=12658628; DOI=10.1002/mrd.10292;  
 RA Ishiwata H., Katsuma S., Kizaki K., Patel O.V., Nakano H.,  
 RA Takahashi T., Inai K., Hirasawa A., Shiojima S., Ikawa H., Suzuki Y.,  
 RA Tsujimoto G., Izaike Y., Todoroki J., Hashizume K.;  
 RT "Characterization of gene expression profiles in early bovine  
 pregnancy using a custom cDNA microarray.";  
 RL Mol. Reprod. Dev. 65:9-18(2003).  
 CC -!- FUNCTION: Actins are highly conserved proteins that are involved  
 in various types of cell motility and are ubiquitously expressed  
 in all eukaryotic cells (By similarity).  
 CC -!- SIMILARITY: Belongs to the actin family.

EMBL; AB098974; BAC56464.1; -.  
 HSSP; P02568; 1P8Z.  
 DR GO; GO:0005884; C:actin filament; IEA.  
 DR GO; GO:0003774; F:motor activity; IEA.  
 DR GO; GO:0005200; F:structural constituent of cytoskeleton; IEA.  
 InterPro; IPR004001; Actin.  
 InterPro; IPR004000; Actin\_like.  
 Pfam; PF00022; Actin; 1.  
 PRINTS; PR00190; ACTIN.  
 DR PROSITE; PS00406; ACTINS\_1; 1.  
 DR PROSITE; PS01132; ACTINS\_ACT\_LIKE; 1.  
 KW Structural protein.

FT NON\_TER 1  
 FT NON\_TER 100  
 SQ SEQUENCE 100 AA; 11601 MW; 0E1467471D0019D CRC64;

Query Match 100.0%; Score 76; DB 2; Length 100;  
 Best Local Similarity 100.0%; Pred. No. 2.6e-06;  
 Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 DEAQSKRGILTLYP 15  
 |||||  
 Db 19 DEAQSKRGILTLYP 33

RESULT 14  
 Q86SB7 PRELIMINARY; PRT; 104 AA.  
 ID Q86SB7  
 AC Q86SB7  
 DT 01-JUN-2003 (TrEMBLrel. 24, Created)  
 DT 01-JUN-2003 (TrEMBLrel. 24, Last sequence update)  
 DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)  
 Actin (Fragment).  
 GN Name=ACT;  
 OS Gryllus bimaculatus (Two-spotted cricket).  
 OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;  
 OC Neoptera; Orthopteroidea; Orthoptera; Ensifera; Gryllidae; Gryllinae;  
 OC Gryllus.

OX NCBI\_TaxID=6999;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA Zhang H., Shimmyo Y., Hirose A., Mito T., Inoue Y., Ohuchi H.,  
 RA Loukeris T.G., Noji S.;  
 RT "Extrachromosomal transposition of the transposable element Minos  
 occurs in embryos of the cricket Gryllus bimaculatus.";

Query Match 100.0%; Score 76; DB 2; Length 100;  
 Best Local Similarity 100.0%; Pred. No. 2.6e-06;  
 Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 DEAQSKRGILTLYP 15  
 |||||  
 Db 19 DEAQSKRGILTLYP 33

```
RL Dev. Growth Differ. 0:0-0(2002).
CC -!- FUNCTION: Actins are highly conserved proteins that are involved
CC in various types of cell motility and are ubiquitously expressed
CC in all eukaryotic cells (By similarity).
CC -!- SIMILARITY: Belongs to the actin family.
DR EMBL; AB087882; BAC55093.1; -.
DR HSSP; P02577; 1NMI.
DR GO; GO:0005884; C:actin filament; IEA.
DR GO; GO:0003774; F:motor activity; IEA.
DR GO; GO:0005200; F:structural constituent of cytoskeleton; IEA.
DR InterPro; IPR004001; Actin.
DR InterPro; IPR004000; Actin_like.
DR Pfam; PF00022; Actin; 1.
DR PRINTS; PR00190; ACTIN.
DR PROSITE; PS00406; ACTINS_1; 1.
KW Structural protein.
FT NON_TER 104
SQ SEQUENCE 104 AA; 11511 MW; 332777A2B08669B0 CRC64;

Query Match 100.0%; Score 76; DB 2; Length 104;
Best Local Similarity 100.0%; Pred.No. 2.7e-06; Mismatches 0; Indels 0; Gaps 0;
Matches 15; Conservative 0;

QY 1 DEAQSKRGILTLYP 15
DB 57 DEAQSKRGILTLYP 71
|||||

RESULT 15
O18548
ID O18548 PRELIMINARY; PRT; 108 AA.
AC O18548;
DT 01-JAN-1998 (TrEMBLrel. 05, Created)
DT 01-JAN-1998 (TrEMBLrel. 05, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE Cytoplasmic actin Lvc4 (Fragment).
GN Name=Lvc4;
OS Lytechinus variegatus (Sea urchin).
OC Eukaryota; Metazoa; Echinodermata; Eleutherozoa; Echinozoa;
OC Echinoidea; Euechinoidea; Echinacea; Temnopleuroidea; Toxopneustidae;
OC Lytechinus.
OX NCBI_TaxID=7654;
RN [1]
RP SEQUENCE FROM N.A.
RA MEDLINE=97333945; PubMed=9190067;
RX Kiseinger J.C., Hahn J.H., Raff R.A.;
RT "Rapid evolution in a conserved gene family. Evolution of the actin
RT gene family in the sea urchin genus Helicodaris and related genera.";
RL Mol. Biol. Evol. 14:654-665(1997).
CC -!- FUNCTION: Actins are highly conserved proteins that are involved
CC in various types of cell motility and are ubiquitously expressed
CC in all eukaryotic cells (By similarity).
CC -!- SIMILARITY: Belongs to the actin family.
DR EMBL; U82542; AAB66299.1; -.
DR HSSP; P10983; 1D4X.
DR GO; GO:0005884; C:actin filament; IEA.
DR GO; GO:0003774; F:motor activity; IEA.
DR GO; GO:0005200; F:structural constituent of cytoskeleton; IEA.
DR InterPro; IPR004001; Actin.
DR InterPro; IPR004000; Actin_like.
DR Pfam; PF00022; Actin; 1.
DR PRINTS; PR00190; ACTIN.
DR PROSITE; PS00406; ACTINS_1; 1.
DR PROSITE; PS01132; ACTINS_ACT_LIKE; 1.
KW Structural protein.
FT NON_TER 108
SQ SEQUENCE 108 AA; 12094 MW; F670CB8A6CBD3AEA CRC64;

Query Match 100.0%; Score 76; DB 2; Length 108;
Best Local Similarity 100.0%; Pred.No. 2.8e-06; Mismatches 0; Indels 0; Gaps 0;
Matches 15; Conservative 0;
```

```
QY 1 DEAQSKRGILTLYP 15
DB 43 DEAQSKRGILTLYP 57
|||||

Search completed: April 8, 2005, 12:03:12
Job time : 58.8667 secs
```



GenCore version 5.1.6  
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: April 8, 2005, 09:58:56 ; Search time 62.4 seconds  
(without alignments)  
92.971 Million cell updates/sec

Title: US-09-423-351C-10  
Perfect score: 71  
Sequence: 1 IQAVLSLYASGRITG 15

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 2105692 seqs, 386760381 residues

Total number of hits satisfying chosen parameters: 2105692

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : A\_Geneseq\_16Dec04:\*

- 1: Geneseqp1980s:\*
- 2: Geneseqp1990s:\*
- 3: Geneseqp2000s:\*
- 4: Geneseqp2001s:\*
- 5: Geneseqp2002s:\*
- 6: Geneseqp2003as:\*
- 7: Geneseqp2003bs:\*
- 8: Geneseqp2004s:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	71	100.0	15	2 AAW92536	Beta-acti
2	71	100.0	97	4 AAM13687	Peptide #
3	71	100.0	97	4 ABB32618	Peptide #
4	71	100.0	97	4 AAM26087	Peptide #
5	71	100.0	97	4 ABB27467	Human pep
6	71	100.0	97	4 ABB18116	Protein #
7	71	100.0	97	4 AAM65825	Human bon
8	71	100.0	97	4 AAM53448	Human bra
9	71	100.0	97	4 ABB47471	Human liv
10	71	100.0	97	4 AAM01437	Peptide #
11	71	100.0	97	5 ABB35459	Human pep
12	71	100.0	137	4 AAU33076	Novel hum
13	71	100.0	150	4 ABB15101	Novel hum
14	71	100.0	157	4 ABB66853	Drosophil
15	71	100.0	166	4 ABB26869	Novel hum
16	71	100.0	168	5 ABB42912	Human ova
17	71	100.0	183	5 ABB42300	Human ova
18	71	100.0	185	6 ABB70840	Human ova
19	71	100.0	196	6 ABB70816	Human adi
20	71	100.0	204	5 ABB43016	Human ova
21	71	100.0	211	7 ABB73899	Human clone
22	71	100.0	239	7 ABB31206	Human dia
23	71	100.0	257	3 AAG37992	Arabidops
24	71	100.0	257	3 AAG20916	Arabidops
25	71	100.0	274	6 ABB70549	Human adi

26	71	100.0	276	3 AAG50947	Arabidops
27	71	100.0	277	3 AAG05115	Arabidops
28	71	100.0	294	3 AAG20915	Arabidops
29	71	100.0	294	3 AAG37991	Arabidops
30	71	100.0	294	7 ADL63068	Human apo
31	71	100.0	295	3 AAG20914	Arabidops
32	71	100.0	295	3 AAG37990	Arabidops
33	71	100.0	304	3 AAG50946	Arabidops
34	71	100.0	305	3 AAG05114	Arabidops
35	71	100.0	308	3 AAG37210	Arabidops
36	71	100.0	311	3 AAG37209	Arabidops
37	71	100.0	321	3 AAG50945	Arabidops
38	71	100.0	322	3 AAG05113	Arabidops
39	71	100.0	332	3 AAG06448	Arabidops
40	71	100.0	332	3 AAG50873	Arabidops
41	71	100.0	332	3 AAG07390	Arabidops
42	71	100.0	332	3 AAG41511	Arabidops
43	71	100.0	332	3 AAG37989	Arabidops
44	71	100.0	332	3 AAG37317	Arabidops
45	71	100.0	332	3 AAG44697	Zea may

## ALIGNMENTS

RESULT 1  
AAW92536  
ID AAW92536 standard; peptide; 15 AA.  
XX AC AAW92536;  
XX 26-APR-1999 (first entry)  
DT 26-APR-1999 (first entry)  
DE Beta-actin reference peptide substrate #10.  
XX Peptide substrate; CCT; eukaryotic type II chaperonin complex; cyclin;  
KW binding agent; substrate-binding site; SBS; substrate folding; actin;  
KW tubulin; treatment; cancer; anticancer drug; viral infection; screening;  
KW reduced toxicity.  
XX Synthetic.  
XX WO9853322-A1.  
XX 26-NOV-1998.  
PD 26-NOV-1998.  
PF 22-MAY-1998; 98WO-GB001485.  
XX 23-MAY-1997; 97GB-00010762.  
XX (CANC-) INST CANCER RES ROYAL CANCER HOSPITAL.  
XX Willison K, Hynes G, Liou AK;  
WPI; 1999-070162/06.  
XX Identifying specific binding agents for substrate binding site in CCT  
PT chaperonin complex - also new peptide binding agents and their mimetics,  
PT and peptides containing a specific CCT binding site, used for treating  
cancer.  
PS Disclosure; Fig 10; 97pp; English.  
XX This invention describes a method which uses the CCT (eukaryotic type II  
CC chaperonin) complex or part of it, for identifying a binding agent that  
CC can occupy a substrate-binding site (SBS) on the CCT complex. By binding  
CC to the CCT complex, the binding agents block an SBS so that biological  
CC activity of the CCT complex is affected, particularly its ability to fold  
CC substrates such as actin, tubulin and cyclin. The binding agents are  
CC useful for treatment of cancer, particularly when used in combination  
CC with an anticancer drug, or viral infections. Nucleic acid fragments are  
CC used to screen for agents, e.g. binding agents that modulate interaction  
CC between the CCT complex and a protein that is to be folded. The binding

CC agents may target cells that are actively synthesizing tubulin etc.  
 CC (unlike known microtubule-stabilising agents that affect all cells), so  
 CC should have reduced toxicity for normal cells. AAM9527-W92541 are  
 CC peptide substrates used in the method of the invention  
 XX  
 SQ Sequence 15 AA;

Query Match 100.0%; Score 71; DB 2; Length 15;  
 Best Local Similarity 100.0%; Pred. No. 8.2e-06;  
 Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 1 IQAVLSLYASGRRTTG 15  
 |||||  
 Db 1 IQAVLSLYASGRRTTG 15  
 |||||

RESULT 2  
 AAM13687  
 ID AAM13687 standard; protein; 97 AA.  
 XX  
 AC AAM13687;  
 XX  
 XX  
 DT 12-OCT-2001 (first entry)  
 XX  
 DE Peptide #121 encoded by probe for measuring cervical gene expression.  
 XX  
 KW Probe; human; microarray; gene expression; cervical epithelial cell;  
 KW cervical cancer.  
 XX  
 OS Homo sapiens.  
 XX  
 PN WO200157278-A2.  
 XX  
 PD 09-AUG-2001.  
 XX  
 PF 30-JAN-2001; 2001WO-US000670.  
 XX  
 PR 04-FEB-2000; 2000US-0180312P.  
 PR 26-MAY-2000; 2000US-0207456P.  
 PR 30-JUN-2000; 2000US-00608408.  
 PR 03-AUG-2000; 2000US-00632366.  
 PR 21-SEP-2000; 2000US-0234687P.  
 PR 27-SEP-2000; 2000US-0236359P.  
 PR 04-OCT-2000; 2000GB-00024263.  
 XX  
 PA (MOLE-) MOLECULAR DYNAMICS INC.  
 XX  
 PI Penn SG, Hanzel DK, Chen W, Rank DR;  
 XX  
 XX WPI; 2001-488901/53.

Human genome-derived single exon nucleic acid probes useful for analyzing  
 gene expression in human cervical epithelial cells.

Claim 27; SEQ ID NO 18513; 487pp; English.

The present invention relates to human single exon nucleic acid probes  
 (SENPs: see AAI10068-AAI28459). The present sequence is a peptide encoded  
 by one such probe. The SENPs are derived from human HeLa cells. The SENPs  
 can be used to produce a single exon microarray, which can be used for  
 measuring human gene expression in a sample derived from human cervical  
 epithelial cells. By measuring gene expression, the probes are therefore  
 useful in grading and/or staging of diseases of the cervix, notably  
 cervical cancer. Note: The sequence data for this patent did not form  
 part of the printed specification, but was obtained in electronic format  
 directly from WIPO at ftp.wipo.int/pub/published\_pct\_sequences

Sequence 97 AA;

Query Match 100.0%; Score 71; DB 4; Length 97;  
 Best Local Similarity 100.0%; Pred. No. 7.6e-05;  
 Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 1 IQAVLSLYASGRRTTG 15  
 |||||  
 Db 15 IQAVLSLYASGRRTTG 29  
 |||||

RESULT 3  
 ABB32618  
 ID ABB32618 standard; peptide; 97 AA.  
 XX  
 AC ABB32618;  
 XX  
 DT 04-FEB-2002 (first entry)  
 XX  
 DE Peptide #124 encoded by human foetal liver single exon probe.  
 XX  
 KW Human; foetal liver; gene expression; single exon nucleic acid probe.  
 XX  
 OS Homo sapiens.  
 XX  
 PN WO200157277-A2.  
 XX  
 PD 09-AUG-2001.  
 XX  
 PF 30-JAN-2001; 2001WO-US000669.  
 XX  
 PR 04-FEB-2000; 2000US-0180312P.  
 PR 26-MAY-2000; 2000US-0207456P.  
 PR 30-JUN-2000; 2000US-00608408.  
 PR 03-AUG-2000; 2000US-00632366.  
 PR 21-SEP-2000; 2000US-0234687P.  
 PR 27-SEP-2000; 2000US-0236359P.  
 PR 04-OCT-2000; 2000GB-00024263.  
 XX  
 PA (MOLE-) MOLECULAR DYNAMICS INC.  
 XX  
 PI Penn SG, Hanzel DK, Chen W, Rank DR;  
 XX  
 XX WPI; 2001-483447/52.

Human genome-derived single exon nucleic acid probes useful for analyzing  
 gene expression in human foetal liver.

Claim 27; SEQ ID NO 25253; 639pp + Sequence Listing; English.

The invention relates to a single exon nucleic acid probe for measuring  
 human gene expression in a sample derived from human foetal liver. The  
 single exon nucleic acid probes may be used for predicting, measuring and  
 displaying gene expression in samples derived from human foetal liver. The  
 present sequence is a peptide encoded by a single exon nucleic acid probe  
 of the invention. Note: The sequence data for this patent did not form  
 part of the printed specification, but was obtained in electronic format  
 directly from WIPO at ftp.wipo.int/pub/published\_pct\_sequences

Sequence 97 AA;

Query Match 100.0%; Score 71; DB 4; Length 97;  
 Best Local Similarity 100.0%; Pred. No. 7.6e-05;  
 Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 1 IQAVLSLYASGRRTTG 15  
 |||||  
 Db 15 IQAVLSLYASGRRTTG 29  
 |||||

RESULT 4  
 AAM26087  
 ID AAM26087 standard; protein; 97 AA.  
 XX  
 AC AAM26087;  
 XX  
 DT 17-OCT-2001 (first entry)  
 XX  
 DE Peptide #124 encoded by probe for measuring placental gene expression.

XX Probe; microarray; human; placenta; antenatal diagnosis;  
KW genetic disorder.

XX Homo sapiens.

PN WO200157272-A2.

PD 09-AUG-2001.

XX 30-JAN-2001; 2001WO-US000663.

XX 04-FEB-2000; 2000US-0180312P.

PR 26-MAY-2000; 2000US-0207456P.

PR 30-JUN-2000; 2000US-00608408.

PR 03-AUG-2000; 2000US-00632366.

PR 21-SEP-2000; 2000US-0234687P.

PR 27-SEP-2000; 2000US-0236359P.

PR 04-OCT-2000; 2000GB-00024263.

XX (MOLE-) MOLECULAR DYNAMICS INC.

XX Penn SG, Hanzel DK, Chen W, Rank DR;

XX WPI; 2001-488897/53.

XX Human genome-derived single exon nucleic acid probes useful for analyzing

XX gene expression in human placenta.

XX Claim 27; SEQ ID NO 26356; 654pp; English.

XX The present invention relates to single exon nucleic acid probes (SENP;  
XX see AA131315-A157546). The present sequence is a peptide encoded by one  
XX such probe. The probes are useful for producing a microarray for  
XX predicting, measuring and displaying gene expression in samples derived  
XX from human placenta. The probes are useful for antenatal diagnosis of  
XX human genetic disorders

XX Sequence 97 AA;

Query Match 100.0%; Score 71; DB 4; Length 97;

Best Local Similarity 100.0%; Pred. No. 7.6e-05;

Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 IQAVLSLYASGRRTTG 15

DB 15 IQAVLSLYASGRRTTG 29

RESULT 5

ABB27467

ID ABB27467 standard; peptide; 97 AA.

XX AC ABB27467;

XX 01-FEB-2002 (first entry)

XX Human peptide #118 encoded by breast cell single exon nucleic acid probe.

XX Human; microarray; single exon probe; gene expression; breast; disease;

XX cancer.

XX Homo sapiens.

XX WO200157271-A2.

XX 09-AUG-2001.

XX 30-JAN-2001; 2001WO-US000662.

XX 04-FEB-2000; 2000US-0180312P.

PR 26-MAY-2000; 2000US-0207456P.

PR 30-JUN-2000; 2000US-00608408.

PR 03-AUG-2000; 2000US-00632366.

PR 21-SEP-2000; 2000US-0234687P.

PR 27-SEP-2000; 2000US-0236359P.

PR 04-OCT-2000; 2000GB-00024263.

XX (MOLE-) MOLECULAR DYNAMICS INC.

XX Penn SG, Hanzel DK, Chen W, Rank DR;

XX WPI; 2001-496933/54.

XX New spatially-addressable set of single exon nucleic acid probes, useful  
XX for measuring gene expression in sample derived from human breast,  
XX comprises number of single exon nucleic acid probes.

XX Claim 27; SEQ ID NO 10435; 327pp + Sequence Listing; English.

XX The invention relates to a spatially-addressable set of single exon  
XX nucleic acid probes for measuring gene expression in a sample derived  
XX from human breast and BT 474 cells. The method involves contacting the  
XX probes with a collection of detectably labelled nucleic acids derived  
XX from mRNA of human breast, and then measuring the label bound to each  
XX probe of the microarray. The probes are useful for verifying the  
XX expression of regions of genomic DNA predicted to encode proteins. They  
XX are useful for gene discovery, and for determining predisposition and/or  
XX assessing breast disease. Gene expression analysis is useful for  
XX this invention presents a far greater diversity of probes for measuring  
XX gene expression, with far less bias than expressed sequence tag  
XX microarrays. The method is suitable for rapid production of functional  
XX information from genomic sequence. The present sequence is a peptide  
XX encoded by a single exon nucleic acid probe of the invention. Note: The  
XX sequence data for this patent did not form part of the printed  
XX specification, but was obtained in electronic format directly from WIPO  
XX at ftp.wipo.int/pub/published\_pct\_sequences

XX Sequence 97 AA;

Query Match 100.0%; Score 71; DB 4; Length 97;

Best Local Similarity 100.0%; Pred. No. 7.6e-05;

Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 IQAVLSLYASGRRTTG 15

DB 15 IQAVLSLYASGRRTTG 29

RESULT 6

ABB18116

ID ABB18116 standard; protein; 97 AA.

XX AC ABB18116;

XX 23-JAN-2002 (first entry)

XX Protein #115 encoded by probe for measuring heart cell gene expression.

XX Human; gene expression; heart; microarray; vascular system;  
XX cardiovascular disease; hypertension; cardiac arrhythmia;  
XX congenital heart disease.

XX Homo sapiens.

XX WO200157274-A2.

XX 09-AUG-2001.

XX 30-JAN-2001; 2001WO-US000666.

XX 04-FEB-2000; 2000US-0180312P.

PR 26-MAY-2000; 2000US-0207456P.

PR 30-JUN-2000; 2000US-00608408.

PR 03-AUG-2000; 2000US-00632366.

```
PR 21-SEP-2000; 2000US-0234687P.
PR 27-SEP-2000; 2000US-0236359P.
PR 04-OCT-2000; 2000GB-00024263.
XX
XX (MOLE-) MOLECULAR DYNAMICS INC.
XX
XX Penn SG, Hanzel DK, Chen W, Rank DR;
XX WPI; 2001-488899/53.
XX
XX Single exon nucleic acid probes for analyzing gene expression in human
XX hearts.
XX
XX Claim 15; SEQ ID NO 19886; 530pp; English.
XX
XX The present invention relates to single exon nucleic acid probes for
XX measuring human gene expression in a sample derived from human heart (see
XX AB421535-ABA41305). The present sequence is a protein encoded by one such
XX probe. The probes may be used for predicting, measuring and displaying
XX gene expression in samples derived from the human heart via microarrays.
XX By measuring gene expression, the probes are useful for predicting,
XX diagnosing, grading, staging, monitoring and prognosing diseases of the
XX human heart and vascular system e.g. cardiovascular disease,
XX hypertension, cardiac arrhythmias and congenital heart disease. Note: The
XX sequence data for this patent did not form part of the printed
XX specification, but was obtained in electronic format directly from WIPO
XX at ftp.wipo.int/pub/published_pct_sequences
XX
XX Sequence 97 AA;

Query Match 100.0%; Score 71; DB 4; Length 97;
Best Local Similarity 100.0%; Pred. No. 7.6e-05;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 IQAVLSLYASGRITG 15
Db 15 IQAVLSLYASGRITG 29

RESULT 7
AAM65825
ID AAM65825 standard; protein; 97 AA.
XX
XX AAM65825;
XX
XX 06-NOV-2001 (first entry)
XX
XX Human bone marrow expressed probe encoded protein SEQ ID NO: 26131.
XX
XX Human; bone marrow expressed exon; gene expression analysis; probe;
XX microarray; cancer; leukaemia; lymphoma; myeloma.
XX
XX Homo sapiens.
XX
XX WO200157276-A2.
XX
XX 09-AUG-2001.
XX
XX 30-JAN-2001; 2001WO-US000668.
XX
XX 04-FEB-2000; 2000US-0180312P.
XX 26-MAY-2000; 2000US-0207456P.
XX 30-JUN-2000; 2000US-00608408.
XX 03-AUG-2000; 2000US-00632366.
XX 21-SEP-2000; 2000US-0234687P.
XX 27-SEP-2000; 2000US-0236359P.
XX 04-OCT-2000; 2000GB-00024263.
XX
XX (MOLE-) MOLECULAR DYNAMICS INC.
XX
XX Penn SG, Hanzel DK, Chen W, Rank DR;
XX WPI; 2001-488900/53.

Query Match 100.0%; Score 71; DB 4; Length 97;
Best Local Similarity 100.0%; Pred. No. 7.6e-05;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 IQAVLSLYASGRITG 15
Db 15 IQAVLSLYASGRITG 29

RESULT 8
AAM53448
ID AAM53448 standard; protein; 97 AA.
XX
XX AAM53448;
XX
XX 05-NOV-2001 (first entry)
XX
XX Human brain expressed single exon probe encoded protein SEQ ID NO: 25553.
XX
XX Human; brain expressed exon; gene expression analysis; probe; microarray;
XX Alzheimer's disease; multiple sclerosis; schizophrenia; epilepsy; cancer.
XX
XX Homo sapiens.
XX
XX WO200157275-A2.
XX
XX 09-AUG-2001.
XX
XX 30-JAN-2001; 2001WO-US000667.
XX
XX 04-FEB-2000; 2000US-0180312P.
XX 26-MAY-2000; 2000US-0207456P.
XX 30-JUN-2000; 2000US-00608408.
XX 03-AUG-2000; 2000US-00632366.
XX 21-SEP-2000; 2000US-0234687P.
XX 27-SEP-2000; 2000US-0236359P.
XX 04-OCT-2000; 2000GB-00024263.
XX
XX (MOLE-) MOLECULAR DYNAMICS INC.
XX
XX Penn SG, Hanzel DK, Chen W, Rank DR;
XX WPI; 2001-483446/52.
XX
XX Single exon nucleic acid probes for analyzing gene expression in human
XX brains.
XX
XX Example 4; SEQ ID NO 25553; 650pp + Sequence Listing; English.
XX
XX The present invention provides a number of single exon nucleic acid
XX probes which are derived from genomic sequences expressed in the human
XX brain. They can be used to measure gene expression in brain cell samples,
XX which may enable the diagnosis and improved treatment of nervous system
XX diseases such as Alzheimer's disease, multiple sclerosis, schizophrenia,
XX epilepsy and cancers. The present sequence is a protein encoded by one of
XX the probes of the invention
XX
XX Sequence 97 AA;
```

Query Match 100.0%; Score 71; DB 4; Length 97;  
Best Local Similarity 100.0%; Pred. No. 7.6e-05;  
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 IQAVLSLYASGRRTTG 15  
DB 15 IQAVLSLYASGRRTTG 29

RESULT 9  
ABG47471  
ID ABG47471 standard; peptide; 97 AA.

XX AC ABG47471;  
XX DT 25-FEB-2003 (first entry)  
XX DE Human liver peptide, SEQ ID No 26119.  
XX KW Human; liver; cirrhosis; hyperlipoproteinaemia; hyperlipidaemia;  
XX KW hypercholesterolaemia; coronary heart disease.

XX OS Homo sapiens.  
XX PN WO200157273-A2.  
XX PD 09-AUG-2001.  
XX PF 30-JAN-2001; 2001WO-US000664.  
XX PR 04-FEB-2000; 2000US-0180312P.  
XX PR 26-MAY-2000; 2000US-0207456P.  
XX PR 30-JUN-2000; 2000US-00608408.  
XX PR 03-AUG-2000; 2000US-00632366.  
XX PR 21-SEP-2000; 2000US-0234687P.  
XX PR 27-SEP-2000; 2000US-0236359P.  
XX PR 04-OCT-2000; 2000GB-00024263.  
XX PA (MOLE-) MOLECULAR DYNAMICS INC.

XX PI Penn SG, Hanzel DK, Chen W, Rank DR;  
XX WPI; 2001-488898/53.

XX DR Human genome-derived single exon nucleic acid probes useful for analyzing  
XX PT gene expression in human adult liver.

XX PS Claim 27; SEQ ID NO 26119; 658pp; English.

XX CC The invention relates to a single exon nucleic acid probe (SENP) (I) for  
XX CC measuring human gene expression in a sample derived from human adult  
XX CC liver, comprising one of 13109 defined nucleotide sequences given in the  
XX CC specification (or complements/ fragments). The probe hybridises at high  
XX CC stringency to a nucleic acid molecule expressed in the human adult liver.  
XX CC (I) may be used for predicting, measuring and displaying gene expression  
XX CC in samples derived from human adult liver. The genes identified may be  
XX CC involved in genetic liver diseases such as cirrhosis,  
XX CC hyperlipoproteinaemia, hyperlipidaemia and hypercholesterolaemia which is  
XX CC associated with coronary heart disease. ABG47348-ABG5930 represent human  
XX CC liver single exon encoded peptides of the invention. Note: The sequence  
XX CC information for this patent does not appear in the printed specification  
XX CC but was obtained in electronic format directly from WIPO at  
XX CC ftp.wipo.int/pub/published\_pct\_sequences

XX SQ Sequence 97 AA;

Query Match 100.0%; Score 71; DB 4; Length 97;  
Best Local Similarity 100.0%; Pred. No. 7.6e-05;  
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 IQAVLSLYASGRRTTG 15  
DB 15 IQAVLSLYASGRRTTG 29

RESULT 11  
ABG35459  
ID ABG35459 standard; peptide; 97 AA.

XX AC ABG35459;

DB 15 IQAVLSLYASGRRTTG 29  
RESULT 10  
AAW01437  
ID AAW01437 standard; protein; 97 AA.  
XX AC AAW01437;  
XX DT 09-OCT-2001 (first entry)

XX DE Peptide #119 encoded by probe for measuring human breast gene expression.  
XX KW Probe; human; breast disease; breast cancer; development disorder;  
XX KW inflammatory disease; proliferative breast disease; non-carcinoma tumour.

XX OS Homo sapiens.

XX PN WO200157270-A2.

XX PD 09-AUG-2001.

XX PF 29-JAN-2001; 2001WO-US000661.

XX PR 04-FEB-2000; 2000US-0180312P.

XX PR 26-MAY-2000; 2000US-0207456P.

XX PR 30-JUN-2000; 2000US-00608408.

XX PR 03-AUG-2000; 2000US-00632366.

XX PR 21-SEP-2000; 2000US-0234687P.

XX PR 27-SEP-2000; 2000US-0236359P.

XX PR 04-OCT-2000; 2000GB-00024263.

XX PA (MOLE-) MOLECULAR DYNAMICS INC.

XX PI Penn SG, Hanzel DK, Chen W, Rank DR;

XX WPI; 2001-476286/51.

XX DR Novel single exon nucleic acid probe used to measuring gene expression in  
XX PT a human breast.

XX PS Claim 27; SEQ ID NO 10177; 322pp; English.

XX CC The present invention relates to novel single exon nucleic acid probes  
XX CC (see AAI00010-AA110067). The present sequence is a peptide encoded by one  
XX CC such probe. The probes are useful for measuring human gene expression in  
XX CC a human breast sample, where the probe hybridises at high stringency to a  
XX CC nucleic acid expressed in the human breast. The probes are useful for  
XX CC predicting, diagnosing, grading, staging, monitoring and prognosing  
XX CC diseases of the human breast, particularly those diseases with polygenic  
XX CC aetiology. The diseases include: breast cancer, disorders of development,  
XX CC inflammatory diseases of the breast, fibrocystic changes, proliferative  
XX CC breast disease and non-carcinoma tumours. Note: The sequence data for  
XX CC this patent did not form part of the printed specification, but was  
XX CC obtained in electronic format directly from WIPO at  
XX CC ftp.wipo.int/pub/published\_pct\_sequences

XX SQ Sequence 97 AA;

Query Match 100.0%; Score 71; DB 4; Length 97;  
Best Local Similarity 100.0%; Pred. No. 7.6e-05;  
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 IQAVLSLYASGRRTTG 15  
DB 15 IQAVLSLYASGRRTTG 29

CC histiocytosis, lymphangi leiomyomatosis, pulmonary alveolar proteinosis,  
CC Karagazer syndrome, fibrocystic pulmonary dysplasia, primary ciliary  
CC dyskinesia, pulmonary hypertension and hyaline membrane disease. The  
CC present sequence is a peptide/protein encoded by a single exon probe of  
CC the invention. Note: The sequence data for this patent did not form part  
CC of the printed specification, but was obtained in electronic format  
CC directly from WIPO at [ftp.wipo.int/pub/published\\_pct\\_sequences](http://ftp.wipo.int/pub/published_pct_sequences)  
XX  
XX  
SQ Sequence 97 AA;  
  
Query Match 100.0%; Score 71; DB 5; Length 97;  
Best Local Similarity 100.0%; Pred. No. 7.6e-05;  
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
  
QY 1 IQAVLSLYASGRRTG 15  
||| ||||| |||||  
DB 15 IQAVLSLYASGRRTG 29  
  
RESULT 12  
AAU33076  
ID AAU33076 standard; protein; 137 AA.  
XX  
XX AAU33076;  
XX  
XX 18-DEC-2001 (first entry)  
XX  
XX Novel human secreted protein #3567.  
XX  
XX Human; vaccination; gene therapy; nutritional supplement;  
KW stem cell proliferation; haematopoiesis; nerve tissue regeneration;  
KW immune suppression; immune stimulation; anti-inflammatory; leukaemia.  
XX  
XX Homo sapiens.  
XX  
XX WO200179449-A2.  
XX  
XX 25-OCT-2001.  
XX  
XX 16-APR-2001; 2001WO-US008656.  
XX  
XX 18-APR-2000; 2000US-00552929.  
PR  
PR 26-JAN-2001; 2001US-00770160.  
XX  
XX (HYSE-) HYSEQ INC.  
XX  
XX  
XX Tang YT, Liu C, Drmanac RT;  
PI  
XX  
XX WPI; 2001-611725/70.  
DR  
XX  
XX Nucleic acids encoding a range of human polypeptides, useful in genetic  
PT vaccination, testing and therapy.  
XX  
XX  
XX Claim 20; Page 704-705; 765pp; English.  
XX  
XX The invention relates to novel human secreted polypeptides. The  
CC polypeptides and antibodies to the polypeptides are useful for  
CC determining the presence of or predisposition to a disease associated  
CC with altered levels of polypeptide. The polypeptides are also useful for  
CC identifying agents (agonists and antagonists) that bind to them. Cells  
CC expressing the proteins are useful for identifying a therapeutic agent  
CC for use in treatment of a pathology related to aberrant expression or  
CC physiological interactions of the polypeptide. Vectors comprising the  
CC nucleic acids encoding the polypeptides and cells genetically engineered  
CC to express them are also useful for producing the proteins. The proteins  
CC are useful in genetic vaccination, testing and therapy, and can be used  
CC as nutritional supplements. They may be used to increase stem cell  
CC proliferation; to regulate haematopoiesis; and in bone, cartilage, tendon  
CC and/or nerve tissue growth or regeneration; immune suppression and/or  
CC stimulation; as anti-inflammatory agents; and in treatment of leukaemias.  
CC AAU29510-AAU33304 represent the amino acid sequences of novel human  
XX secreted proteins of the invention  
XX

SQ Sequence 137 AA;  
 Query Match 100.0%; Score 71; DB 4; Length 137;  
 Best Local Similarity 100.0%; Pred. No. 0.00012; Mismatches 0; Indels 0; Gaps 0;  
 Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 IQAVLSLYASGRRTG 15  
 DB 80 IQAVLSLYASGRRTG 94  
 RESULT 13  
 ID ABG15101 standard; protein; 150 AA.  
 AC ABG15101;  
 DT 18-FEB-2002 (first entry)  
 DE Novel human diagnostic protein #15092.  
 KW Human; chromosome mapping; gene mapping; gene therapy; forensic;  
 KW food supplement; medical imaging; diagnostic; genetic disorder.  
 OS Homo sapiens.  
 PN WO200175067-A2.  
 PD 11-OCT-2001.  
 PF 30-MAR-2001; 2001WO-US008631.  
 PR 31-MAR-2000; 2000US-00540217.  
 PR 23-AUG-2000; 2000US-00649167.  
 XX (HYSE-) HYSEQ INC.  
 PI Drmanac RT, Liu C, Tang YT;  
 DR WPI: 2001-639362/73.  
 DR N-PSDB; AAS79288.  
 XX New isolated polynucleotide and encoded polypeptides, useful in  
 PT diagnostics, forensics, gene mapping, identification of mutations  
 PT responsible for genetic disorders or other traits and to assess  
 PT biodiversity.  
 PS Claim 20; SEQ ID NO 45460; 103pp; English.  
 CC The invention relates to isolated polynucleotide (I) and polypeptide (II)  
 CC sequences. (I) is useful as hybridisation probes, polymerase chain  
 CC reaction (PCR) primers, oligomers, and for chromosome and gene mapping,  
 CC and in recombinant production of (II). The polynucleotides are also used  
 CC in diagnostics as expressed sequence tags for identifying expressed  
 CC genes. (II) is useful in gene therapy techniques to restore normal  
 CC activity of (II) or to treat disease states involving (II). (II) is  
 CC useful for generating antibodies against it, detecting or quantitating a  
 CC polypeptide in tissue, as molecular weight markers and as a food  
 CC supplement. (II) and its binding partners are useful in medical imaging  
 CC of sites expressing (II). (I) and (II) are useful for treating disorders  
 CC involving aberrant protein expression or biological activity. The  
 CC polypeptide and polynucleotide sequences have applications in  
 CC diagnostics, forensics, gene mapping, identification of mutations  
 CC responsible for genetic disorders or other traits to assess biodiversity  
 CC and to produce other types of data and products dependent on DNA and  
 CC amino acid sequences. ABG0010-ABG30377 represent novel human diagnostic  
 CC amino acid sequences of the invention. Note: The sequence data for this  
 CC patent did not appear in the printed specification, but was obtained in  
 CC electronic format directly from WIPO at  
 CC ftp.wipo.int/pub/published\_pct\_sequences  
 XX Sequence 150 AA;  
 SQ  
 Query Match 100.0%; Score 71; DB 4; Length 150;  
 Best Local Similarity 100.0%; Pred. No. 0.00013; Mismatches 0; Indels 0; Gaps 0;  
 Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 IQAVLSLYASGRRTG 15  
 DB 98 IQAVLSLYASGRRTG 112  
 RESULT 14  
 ID ABB66853 standard; protein; 157 AA.  
 AC ABB66853;  
 DT 26-MAR-2002 (first entry)  
 DE Drosophila melanogaster polypeptide SEQ ID NO 27351.  
 KW Drosophila; developmental biology; cell signalling; insecticide;  
 KW pharmaceutical.  
 OS Drosophila melanogaster.  
 PN WO200171042-A2.  
 PD 27-SEP-2001.  
 PF 23-MAR-2001; 2001WO-US009231.  
 PR 23-MAR-2000; 2000US-0191637P.  
 PR 11-JUL-2000; 2000US-00614150.  
 XX (PEKE) PE CORP NY.  
 PI Venter JC, Adams M, Li PWD, Myers EW;  
 DR WPI: 2001-656860/75.  
 DR N-PSDB; ABL10956.  
 XX New isolated nucleic acid detection reagent for detecting 1000 or more  
 PT genes from Drosophila and for elucidating cell signaling and cell-cell  
 PT interactions.  
 PS Disclosure; SEQ ID NO 27351; 21pp + Sequence Listing; English.  
 CC The invention relates to an isolated nucleic acid detection reagent  
 CC capable of detecting 1000 or more genes from Drosophila. The invention is  
 CC useful in developmental biology and in elucidating cell signalling and  
 CC cell-cell interactions in higher eukaryotes for the development of  
 CC insecticides, therapeutics and pharmaceutical drugs. The invention  
 CC discloses genomic DNA sequences (ABL16176-ABL30511), expressed DNA  
 CC sequences (ABL01840-ABL16175) and the encoded proteins (ABB57737-  
 CC ABB72072). The sequence data for this patent did not form part of the  
 CC printed specification, but was obtained in electronic format directly  
 CC from WIPO at ftp.wipo.int/pub/published\_pct\_sequences  
 XX Sequence 157 AA;  
 SQ  
 Query Match 100.0%; Score 71; DB 4; Length 157;  
 Best Local Similarity 100.0%; Pred. No. 0.00014; Mismatches 0; Indels 0; Gaps 0;  
 Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 IQAVLSLYASGRRTG 15  
 DB 137 IQAVLSLYASGRRTG 151  
 RESULT 15  
 ID ABB26869 standard; protein; 166 AA.  
 AC ABB26869;  
 XX ABB26869;  
 SQ

```

XX 18-FEB-2002 (first entry)
XX
XX Novel human diagnostic protein #26860.
XX
XX Human; chromosome mapping; gene mapping; gene therapy; forensic;
XX food supplement; medical imaging; diagnostic; genetic disorder.
XX
XX Homo sapiens.
XX
XX WO200175067-A2.
XX
XX PD 11-OCT-2001.
XX
XX 30-MAR-2001; 2001WO-US008631.
XX
XX 31-MAR-2000; 2000US-00540217.
XX
XX 23-AUG-2000; 2000US-00649167.
XX
XX (HYSE-) HYSEQ INC.
XX
XX Drmanac RT, Liu C, Tang YT;
XX
XX WPI: 2001-639362/73.
XX
XX N-PSDB; AAS91056.
XX
XX New isolated polynucleotide and encoded polypeptides, useful in
XX diagnostics, forensics, gene mapping, identification of mutations
XX responsible for genetic disorders or other traits and to assess
XX biodiversity.
XX
XX Claim 20; SEQ ID NO 57228; 103pp; English.
XX
XX The invention relates to isolated polynucleotide (I) and polypeptide (II)
XX sequences. (I) is useful as hybridisation probes, polymerase chain
XX reaction (PCR) primers, oligomers, and for chromosome and gene mapping,
XX and in recombinant production of (II). The polynucleotides are also used
XX in diagnostics as expressed sequence tags for identifying expressed
XX genes. (I) is useful in gene therapy techniques to restore normal
XX activity of (II) or to treat disease states involving (II). (II) is
XX useful for generating antibodies against it, detecting or quantitating a
XX polypeptide in tissue, as molecular weight markers and as a food
XX supplement. (II) and its binding partners are useful in medical imaging
XX of sites expressing (II). (I) and (II) are useful for treating disorders
XX involving aberrant protein expression or biological activity. The
XX polypeptide and polynucleotide sequences have applications in
XX diagnostics, forensics, gene mapping, identification of mutations
XX and to produce other types of data and products dependent on DNA and
XX amino acid sequences. ABG00010-ABG30377 represent novel human diagnostic
XX amino acid sequences of the invention. Note: The sequence data for this
XX patent did not appear in the printed specification, but was obtained in
XX electronic format directly from WIPO at
XX ftp.wipo.int/pub/published_pct_sequences
XX
XX Sequence 166 AA;
XX
XX Query Match 100.0%; Score 71; DB 4; Length 166;
XX Best Local Similarity 100.0%; Pred. NO. 0.00015;
XX Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
XX
XX QY 1 IQAVLSLYASGRRTG 15
XX
XX DB 147 IQAVLSLYASGRRTG 161
XX
XX Search completed: April 8, 2005, 10:50:48
XX Job time : 62.4 secs

```



GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: April 8, 2005, 09:58:57 ; Search time 17.5333 Seconds  
(without alignments)  
63.863 Million cell updates/sec

Title: US-09-423-351C-10

Perfect score: 71

Sequence: 1 IQAVLSLYASGRRTTG 15

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 513545 seqs, 74649064 residues

Total number of hits satisfying chosen parameters: 513545

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Issued Patents AA:\*

- 1: /cgn2\_6/ptodata/1/1aa/5A COMB.pdp:\*
- 2: /cgn2\_6/ptodata/1/1aa/5B COMB.pdp:\*
- 3: /cgn2\_6/ptodata/1/1aa/6A COMB.pdp:\*
- 4: /cgn2\_6/ptodata/1/1aa/6B COMB.pdp:\*
- 5: /cgn2\_6/ptodata/1/1aa/PCTUS COMB.pdp:\*
- 6: /cgn2\_6/ptodata/1/1aa/backfiles1.pdp:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	71	100.0	336	4	US-09-248-796A-14108 Sequence 14108, A
2	71	100.0	371	1	US-08-261-206A-77 Sequence 77, Appl
3	71	100.0	374	3	US-08-609-236-6 Sequence 6, Appl
4	71	100.0	375	2	US-08-494-151-14 Sequence 14, Appl
5	71	100.0	375	3	US-09-106-217-16 Sequence 16, Appl
6	71	100.0	375	3	US-09-171-337A-7 Sequence 7, Appl
7	71	100.0	375	3	US-09-171-337A-8 Sequence 8, Appl
8	71	100.0	375	4	US-09-631-022-7 Sequence 7, Appl
9	71	100.0	375	4	US-09-631-022-8 Sequence 8, Appl
10	71	100.0	375	4	US-09-976-594-731 Sequence 731, Appl
11	71	100.0	376	1	US-08-588-113-2 Sequence 2, Appl
12	71	100.0	376	4	US-09-949-016-6100 Sequence 6100, Ap
13	71	100.0	377	3	US-09-106-217-2 Sequence 2, Appl
14	71	100.0	377	4	US-09-919-172-33 Sequence 33, Appl
15	71	100.0	377	4	US-09-917-254-53 Sequence 53, Appl
16	71	100.0	386	4	US-09-949-016-7721 Sequence 7721, Ap
17	71	100.0	399	4	US-09-949-016-9424 Sequence 9424, Ap
18	71	100.0	402	4	US-09-949-016-10757 Sequence 10757, A
19	71	100.0	404	4	US-09-949-016-11313 Sequence 11313, A
20	68	95.8	377	4	US-09-248-796A-14109 Sequence 14109, A
21	65	91.5	146	3	US-09-306-446C-5 Sequence 5, Appl
22	65	91.5	362	4	US-09-949-016-7725 Sequence 7725, Ap
23	65	91.5	374	3	US-09-306-446C-2 Sequence 2, Appl
24	65	91.5	376	4	US-09-538-092-1109 Sequence 1109, Ap
25	65	91.5	376	4	US-09-538-092-1110 Sequence 1110, Ap
26	65	91.5	376	4	US-09-949-016-6656 Sequence 6656, Ap
27	65	91.5	376	4	US-09-949-016-8452 Sequence 8452, Ap

28	62	87.3	384	4	US-09-538-092-395 Sequence 395, App
29	51	71.8	394	4	US-09-949-016-6655 Sequence 6655, Ap
30	51	71.8	406	4	US-09-949-016-7396 Sequence 7396, Ap
31	44	62.0	3815	3	US-09-428-517-3 Sequence 3, Appl
32	44	62.0	4150	3	US-09-428-517-2 Sequence 2, Appl
33	39	54.9	888	4	US-09-697-022-4 Sequence 4, Appl
34	39	54.9	888	4	US-09-697-022-5 Sequence 5, Appl
35	39	54.9	6396	4	US-09-410-551B-72 Sequence 72, Appl
36	39	54.9	6396	4	US-09-940-316B-72 Sequence 72, Appl
37	38	53.5	138	4	US-09-107-532A-6906 Sequence 6906, Ap
38	38	53.5	389	4	US-09-270-767-39072 Sequence 39072, A
39	38	53.5	389	4	US-09-270-767-54289 Sequence 54289, A
40	38	53.5	953	3	US-09-245-281-43 Sequence 43, Appl
41	38	53.5	953	4	US-09-207-359B-43 Sequence 43, Appl
42	38	53.5	953	4	US-09-340-620A-43 Sequence 43, Appl
43	38	53.5	953	4	US-09-865-364-43 Sequence 43, Appl
44	38	53.5	966	4	US-09-207-359B-47 Sequence 47, Appl
45	38	53.5	966	4	US-09-865-364-47 Sequence 47, Appl

## ALIGNMENTS

### RESULT 1

US-09-248-796A-14108  
; Sequence 14108, Application US/09248796A  
; Patent No. 6747137  
; GENERAL INFORMATION:  
; APPLICANT: Keith Weinstock et al  
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO CANDIDA ALBICAN  
; FILE REFERENCE: 107196.132  
; CURRENT APPLICATION NUMBER: US/09/248,796A  
; CURRENT FILING DATE: 1999-02-12  
; PRIOR APPLICATION NUMBER: US 60/074,725  
; PRIOR FILING DATE: 1998-02-13  
; PRIOR APPLICATION NUMBER: US 60/096,409  
; PRIOR FILING DATE: 1998-08-13  
; NUMBER OF SEQ ID NOS: 28208  
; SEQ ID NO 14108  
; LENGTH: 336  
; TYPE: PRT  
; ORGANISM: Candida albicans  
US-09-248-796A-14108

Query Match 100.0%; Score 71; DB 4; Length 336;  
Best Local Similarity 100.0%; Pred. No. 7.6e-05;  
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 1 IQAVLSLYASGRRTTG 15

Db 69 IQAVLSLYASGRRTTG 83

### RESULT 2

US-08-261-206A-77  
; Sequence 77, Application US/08261206A  
; Patent No. 5574007  
; GENERAL INFORMATION:  
; APPLICANT: Zushi, Mitichitaka  
; APPLICANT: Gomi, Komakazu  
; APPLICANT: Yamamoto, Shuji  
; APPLICANT: Suzuki, Koji  
; APPLICANT: Matsuda, Akio  
; TITLE OF INVENTION: A Polypeptide Capable of Interacting  
; TITLE OF INVENTION: with Thrombin  
; NUMBER OF SEQUENCES: 80  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Birch, Stewart, Kolasch & Birch  
; STREET: 301 N. Washington St.  
; CITY: Falls Church  
; STATE: Virginia  
; COUNTRY: USA

ZIP: 22046-0747  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.25  
CURRENT APPLICATION DATA: US/08/261,206A  
FILING DATE:  
CLASSIFICATION: 530  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 530  
FILING DATE: 03-AUG-1991  
ATTORNEY/AGENT INFORMATION:  
NAME: Svensson, Leonard R.  
REGISTRATION NUMBER: 30330  
REFERENCE/DOCKET NUMBER: 216-275P  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 703-241-1300  
TELEFAX: 703-241-2848  
TELEX: 248345  
INFORMATION FOR SEQ ID NO: 77:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 371 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-261-206A-77

Query Match 100.0%; Score 71; DB 1; Length 371;  
Best Local Similarity 100.0%; Pred. No. 8.5e-05;  
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 IQAVLSLYASGRRTG 15  
Db 134 IQAVLSLYASGRRTG 148

RESULT 3  
US-08-609-236-6  
Sequence 6, Application US/08609236  
Patent No. 6087398  
GENERAL INFORMATION:  
APPLICANT: Steven R. Goodman  
TITLE OF INVENTION: No. 6087398el Sickie Cell Anemia Treatment  
NUMBER OF SEQUENCES: 6  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: McGregor & Adler, LLP  
STREET: 8011 Candle Lane  
CITY: Houston  
STATE: Texas  
COUNTRY: USA  
ZIP: 77071  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: Apple Macintosh  
OPERATING SYSTEM: Macintosh  
SOFTWARE: Microsoft Word for Macintosh  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/609,236  
FILING DATE: March 1, 1996  
CLASSIFICATION: 514  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 60/002,288  
FILING DATE: August 14, 1995  
CLASSIFICATION: 514  
ATTORNEY/AGENT INFORMATION:  
NAME: Benjamin Aaron Adler, Ph.D.  
REGISTRATION NUMBER: 35,423  
REFERENCE/DOCKET NUMBER: D5807  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 713-777-2321  
TELEFAX: 713-777-6908

INFORMATION FOR SEQ ID NO: 6:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 374  
TYPE: Amino acid  
STRANDEDNESS:  
TOPOLOGY: Linear  
MOLECULE TYPE: Protein  
DESCRIPTION: Protein  
HYPOTHETICAL: No  
ANTI-SENSE: No  
FRAGMENT TYPE:  
ORIGINAL SOURCE:  
STRAIN:  
INDIVIDUAL ISOLATE:  
DEVELOPMENTAL STAGE:  
TISSUE TYPE:  
CELL TYPE:  
CELL LINE:  
US-08-609-236-6

Query Match 100.0%; Score 71; DB 3; Length 374;  
Best Local Similarity 100.0%; Pred. No. 8.6e-05;  
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 IQAVLSLYASGRRTG 15  
Db 135 IQAVLSLYASGRRTG 149

RESULT 4  
US-08-494-151-14  
Sequence 14, Application US/08494151  
Patent No. 5840528  
GENERAL INFORMATION:  
APPLICANT: Van Ooyen, Albert Johannes Joseph  
TITLE OF INVENTION: Transformation of Phaffia rhodozyma  
NUMBER OF SEQUENCES: 14  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Morrison & Foerster  
STREET: 2000 Pennsylvania Ave. N.W., Suite 5500  
CITY: Washington, D.C.  
COUNTRY: USA  
ZIP: 20006-1812  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/494,151  
FILING DATE: 23-JUN-1995  
CLASSIFICATION: 435  
ATTORNEY/AGENT INFORMATION:  
NAME: Murashige, Kate H.  
REGISTRATION NUMBER: 29,959  
REFERENCE/DOCKET NUMBER: 24615-20039.01  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (202) 887-1500  
TELEFAX: (202) 887-0763  
TELEX: 90-4030  
INFORMATION FOR SEQ ID NO: 14:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 375 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-494-151-14

Query Match 100.0%; Score 71; DB 2; Length 375;  
Best Local Similarity 100.0%; Pred. No. 8.6e-05;  
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 IQAVLSLYASGRRTG 15

Db 136 IQAVLSLYASGRRTG 150  
|||||

## RESULT 5

US-09-106-217-16  
; Sequence 16, Application US/09106217  
; Patent No. 6063576  
; GENERAL INFORMATION:  
; APPLICANT: Keating, Mark T.  
; APPLICANT: Olson, Timothy M.  
; TITLE OF INVENTION: Actin Mutations in Dilated  
; TITLE OF INVENTION: Cardiomyopathy, a Heritable Form of Heart Failure  
; NUMBER OF SEQUENCES: 18  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Rothwell, Figg, Ernst & Kurz, P.C.  
; STREET: 555 Thirteenth Street, N.W., Suite 701 East  
; CITY: Washington  
; STATE: DC  
; COUNTRY: U.S.A.  
; ZIP: 20004  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patent in Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/09/106,217  
; FILING DATE:  
; CLASSIFICATION:  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Saxe, Stephen A.  
; REGISTRATION NUMBER: 38,609  
; REFERENCE/DOCKET NUMBER: 2323-125  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 202-783-6040  
; TELEFAX: 202-783-6031  
; INFORMATION FOR SEQ ID NO: 16:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 375 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
US-09-106-217-16

Query Match 100.0%; Score 71; DB 3; Length 375;  
Best Local Similarity 100.0%; Pred. No. 8.6e-05;  
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 IQAVLSLYASGRRTG 15  
|||||

## RESULT 6

US-09-171-337A-7  
; Sequence 7, Application US/09171337A  
; Patent No. 6300095  
; GENERAL INFORMATION:  
; APPLICANT: BARREDO FUENTE, Jose Luis  
; RODRIGUEZ SAIZ, Marta  
; COLLADOS DE LA VIEJA, Alfonso J.  
; MORENO VALLE, Migueu Angel  
; SALTO MALDONADO, Francisco  
; DIEZ GARCIA, Bruno  
; TITLE OF INVENTION: PROMOTERS OF THE GENES GLUTAMATE  
; DESHYDROGENASE, -N-ACETYLHEXOSAMINIDASE  
; AND -ACTIN AND THEIR USE IN FILAMENTOUS  
; FUNGI EXPRESSION, SECRETION AND ANTISENSE  
; NUMBER OF SEQUENCES: 20  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: LADAS & PARRY

STREET: 26 WEST 61 STREET  
CITY: NEW YORK  
STATE: NY  
COUNTRY: USA  
ZIP: 10023  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: 3-1/4" Disk 1.44MB  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: Microsoft Windows for Workgroups 3.11  
; SOFTWARE: WordPerfect 8 for Windows  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/09/171,337A  
; FILING DATE: 14-May-1999  
; CLASSIFICATION: 536  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: PCT/ES98/00056  
; FILING DATE: 5-MAR-1998  
; APPLICATION NUMBER: ES9700482  
; FILING DATE: 5-MAR-1997  
; ATTORNEY/AGENT INFORMATION:  
; NAME: MASS, Clifford J.  
; REGISTRATION NUMBER: 30,086  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 233288  
; INFORMATION FOR SEQ ID NO: 7  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 375 amino acids  
; TYPE: amino acids  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; MOLECULE TYPE: peptide  
; ORIGINAL SOURCE:  
; ORGANISM: Penicillium chrysogenum  
; FEATURE:  
; OTHER INFORMATION: amino acid sequence of the -actin  
; protein with a molecular weight of  
; 41760 Da.  
; SEQUENCE DESCRIPTION: SEQ ID NO: 7  
US-09-171-337A-7

Query Match 100.0%; Score 71; DB 3; Length 375;  
Best Local Similarity 100.0%; Pred. No. 8.6e-05;  
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 IQAVLSLYASGRRTG 15  
|||||

## RESULT 7

US-09-171-337A-8  
; Sequence 8, Application US/09171337A  
; Patent No. 6300095  
; GENERAL INFORMATION:  
; APPLICANT: BARREDO FUENTE, Jose Luis  
; RODRIGUEZ SAIZ, Marta  
; COLLADOS DE LA VIEJA, Alfonso J.  
; MORENO VALLE, Migueu Angel  
; SALTO MALDONADO, Francisco  
; DIEZ GARCIA, Bruno  
; TITLE OF INVENTION: PROMOTERS OF THE GENES GLUTAMATE  
; DESHYDROGENASE, -N-ACETYLHEXOSAMINIDASE  
; AND -ACTIN AND THEIR USE IN FILAMENTOUS  
; FUNGI EXPRESSION, SECRETION AND ANTISENSE  
; NUMBER OF SEQUENCES: 20  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: LADAS & PARRY  
; STREET: 26 WEST 61 STREET  
; CITY: NEW YORK  
; STATE: NY  
; COUNTRY: USA  
; ZIP: 10023



```

US-08-588-113-2
; Sequence 2, Application US/08588113
; Patent No. 5710003
; GENERAL INFORMATION:
; APPLICANT: McHugh, Kirk M.
; TITLE OF INVENTION: DIAGNOSTIC TEST FOR DETERMINING
; TITLE OF INVENTION: MALIGNANCY OF SMOOTH MUSCLE TUMORS
; NUMBER OF SEQUENCES: 4
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Woodcock Washburn Kurtz Mackiewicz &
; ADDRESSEE: No. 5710003ris
; STREET: One Liberty Place - 46th Floor
; CITY: Philadelphia
; STATE: PA
; COUNTRY: USA
; ZIP: 19103
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/588,113
; FILING DATE:
; CLASSIFICATION: 436
; ATTORNEY/AGENT INFORMATION:
; NAME: Ralph, Rebecca L.
; REGISTRATION NUMBER: 35,152
; REFERENCE/DOCKET NUMBER: TJU-1652
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 215-568-3100
; TELEFAX: 215-568-3439
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 376 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-588-113-2

Query Match 100.0%; Score 71; DB 1; Length 376;
Best Local Similarity 100.0%; Pred. No. 8.7e-05;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 IQAVLSLYASGRRTG 15
Db 137 IQAVLSLYASGRRTG 151

RESULT 12
US-09-949-016-6100
; Sequence 6100, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; TITLE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: Fast-SEQ for Windows Version 4.0
; SEQ ID NO 6100
; LENGTH: 376
; TYPE: PRT
; ORGANISM: Human
US-09-949-016-6100

```

RESULT 14  
US-09-919-172-33  
; Sequence 33, Application US/09919172  
; Patent No. 6673545  
; GENERAL INFORMATION:  
; APPLICANT: Paris, Mary  
; APPLICANT: Turner, Christopher M.  
; TITLE OF INVENTION: PROSTATE CANCER MARKERS  
; FILE REFERENCE: PA-0036 US  
; CURRENT APPLICATION NUMBER: US/09/919,172  
; CURRENT FILING DATE: 2001-07-30  
; PRIOR APPLICATION NUMBER: 60/222,469

Search completed: April 8, 2005, 12:07:38  
Job time : 18.5333 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: April 8, 2005, 10:53:18 ; Search time 42.6667 seconds  
(without alignments)  
116.718 Million cell updates/sec

Title: US-09-423-351C-10

Perfect score: 71

Sequence: 1 IQAVLSLYASGRRTTG 15

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1418010 seqs, 331997259 residues

Total number of hits satisfying chosen parameters: 1418010

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

Published Applications AA:\*

- 1: /cgn2\_6/ptodata/1/pubpaa/US07\_PUBCOMB.pep:\*
- 2: /cgn2\_6/ptodata/1/pubpaa/PCT\_NEW\_PUB.pep:\*
- 3: /cgn2\_6/ptodata/1/pubpaa/US05\_NEW\_PUB.pep:\*
- 4: /cgn2\_6/ptodata/1/pubpaa/US06\_PUBCOMB.pep:\*
- 5: /cgn2\_6/ptodata/1/pubpaa/US07\_NEW\_PUB.pep:\*
- 6: /cgn2\_6/ptodata/1/pubpaa/PCTUS\_PUBCOMB.pep:\*
- 7: /cgn2\_6/ptodata/1/pubpaa/US08\_NEW\_PUB.pep:\*
- 8: /cgn2\_6/ptodata/1/pubpaa/US08\_PUBCOMB.pep:\*
- 9: /cgn2\_6/ptodata/1/pubpaa/US09A\_PUBCOMB.pep:\*
- 10: /cgn2\_6/ptodata/1/pubpaa/US09B\_PUBCOMB.pep:\*
- 11: /cgn2\_6/ptodata/1/pubpaa/US09C\_PUBCOMB.pep:\*
- 12: /cgn2\_6/ptodata/1/pubpaa/US09\_NEW\_PUB.pep:\*
- 13: /cgn2\_6/ptodata/1/pubpaa/US10A\_PUBCOMB.pep:\*
- 14: /cgn2\_6/ptodata/1/pubpaa/US10B\_PUBCOMB.pep:\*
- 15: /cgn2\_6/ptodata/1/pubpaa/US10C\_PUBCOMB.pep:\*
- 16: /cgn2\_6/ptodata/1/pubpaa/US10D\_PUBCOMB.pep:\*
- 17: /cgn2\_6/ptodata/1/pubpaa/US10\_NEW\_PUB.pep:\*
- 18: /cgn2\_6/ptodata/1/pubpaa/US11\_NEW\_PUB.pep:\*
- 19: /cgn2\_6/ptodata/1/pubpaa/US60\_NEW\_PUB.pep:\*
- 20: /cgn2\_6/ptodata/1/pubpaa/US60\_PUBCOMB.pep:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match %	Length	ID	Description
1	71	100.0	97	US-09-864-761-33414	Sequence 33414, A
2	71	100.0	168	US-10-264-049-4044	Sequence 4044, Ap
3	71	100.0	180	US-10-437-963-169247	Sequence 169247, Ap
4	71	100.0	183	US-10-264-049-3432	Sequence 3432, Ap
5	71	100.0	197	US-10-424-599-280810	Sequence 280810, Ap
6	71	100.0	204	US-10-264-049-4148	Sequence 4148, Ap
7	71	100.0	208	US-10-424-599-282905	Sequence 282905, Ap
8	71	100.0	219	US-10-002-631C-56	Sequence 56, Appl
9	71	100.0	230	US-10-767-701-42770	Sequence 42770, A
10	71	100.0	236	US-10-424-599-206389	Sequence 206389, Ap
11	71	100.0	237	US-10-424-599-143226	Sequence 143226, Ap
12	71	100.0	248	US-10-424-599-162684	Sequence 162684, Ap
13	71	100.0	249	US-10-767-701-43575	Sequence 43575, A

Sequence 44494, A  
Sequence 69353, A  
Sequence 177530, A  
Sequence 52089, A  
Sequence 57500, A  
Sequence 278431, A  
Sequence 62158, A  
Sequence 59172, A  
Sequence 28333, A  
Sequence 3002, Ap  
Sequence 68, Appl  
Sequence 93, Appl  
Sequence 88, Appl  
Sequence 2436, Ap  
Sequence 5927, Ap  
Sequence 94, Appl  
Sequence 82, Appl  
Sequence 101, Appl  
Sequence 179917, A  
Sequence 202420, A  
Sequence 45677, A  
Sequence 46090, A  
Sequence 33, Appl  
Sequence 52, Appl  
Sequence 64, Appl  
Sequence 162009, A  
Sequence 162685, A  
Sequence 162686, A  
Sequence 22816, A  
Sequence 223492, A  
Sequence 283336, A  
Sequence 66124, A

US-10-425-114-44494  
US-10-425-114-69353  
US-10-424-599-177530  
US-10-425-114-52089  
US-10-425-114-57500  
US-10-424-599-278431  
US-10-425-114-62158  
US-10-425-114-59172  
US-10-424-599-28333  
US-10-128-714-3002  
US-10-322-281-68  
US-10-205-194-93  
US-10-316-253-88  
US-10-369-493-2436  
US-10-369-493-5927  
US-10-205-331-94  
US-10-260-708-82  
US-10-341-434-101  
US-10-437-963-179917  
US-10-437-963-202420  
US-10-767-701-45677  
US-10-767-701-46090  
US-09-913-172-33  
US-10-338-777-52  
US-10-236-031B-64  
US-10-424-599-162009  
US-10-424-599-162685  
US-10-424-599-162686  
US-10-424-599-22816  
US-10-424-599-223492  
US-10-424-599-283336  
US-10-425-114-66124

## ALIGNMENTS

### RESULT 1

US-09-864-761-33414  
; Sequence 33414, Application US/09864761  
; Patent No. US20020048763A1  
; GENERAL INFORMATION:  
; APPLICANT: Penn, Sharon G.  
; APPLICANT: Rank, David R.  
; APPLICANT: Hanzel, David K.  
; APPLICANT: Chen, Wensheng  
; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR  
; FILE REFERENCE: Aemica-X-1  
; CURRENT APPLICATION NUMBER: US/09/864,761  
; CURRENT FILING DATE: 2001-05-23  
; PRIOR APPLICATION NUMBER: US 60/180,312  
; PRIOR FILING DATE: 2000-02-04  
; PRIOR APPLICATION NUMBER: US 60/207,456  
; PRIOR FILING DATE: 2000-05-26  
; PRIOR APPLICATION NUMBER: US 09/632,366  
; PRIOR FILING DATE: 2000-08-03  
; PRIOR APPLICATION NUMBER: GB 24263.6  
; PRIOR FILING DATE: 2000-10-04  
; PRIOR APPLICATION NUMBER: US 60/236,359  
; PRIOR FILING DATE: 2000-09-27  
; PRIOR APPLICATION NUMBER: PCT/US01/00666  
; PRIOR FILING DATE: 2001-01-30  
; PRIOR APPLICATION NUMBER: PCT/US01/00667  
; PRIOR FILING DATE: 2001-01-30  
; PRIOR APPLICATION NUMBER: PCT/US01/00664  
; PRIOR FILING DATE: 2001-01-30  
; PRIOR APPLICATION NUMBER: PCT/US01/00669  
; PRIOR FILING DATE: 2001-01-30  
; PRIOR APPLICATION NUMBER: PCT/US01/00665  
; PRIOR FILING DATE: 2001-01-30  
; PRIOR APPLICATION NUMBER: PCT/US01/00668  
; PRIOR FILING DATE: 2001-01-30

```
; PRIOR APPLICATION NUMBER: PCT/US01/00663
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00662
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00661
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00670
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: US 60/234,687
; PRIOR FILING DATE: 2000-09-21
; PRIOR APPLICATION NUMBER: US 09/608,408
; PRIOR FILING DATE: 2000-06-30
; PRIOR APPLICATION NUMBER: US 09/774,203
; PRIOR FILING DATE: 2001-01-29
; NUMBER OF SEQ ID NOS: 49117
; SOFTWARE: Anomax Sequence Listing Engine vers. 1.1
; SEQ ID NO 33414
; LENGTH: 97
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: MAP TO AC006483.2
; OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 18
; OTHER INFORMATION: EXPRESSED IN HEART, SIGNAL = 48
; OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 81
; OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 28
; OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 24
; OTHER INFORMATION: EXPRESSED IN BT474, SIGNAL = 85
; OTHER INFORMATION: EXPRESSED IN HBL100, SIGNAL = 78
; OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 37
; OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 45
; OTHER INFORMATION: EXPRESSED IN HELA, SIGNAL = 40
; OTHER INFORMATION: EST HUMAN HIT: BE271730.1, EVALUATE 3.00e-50
; OTHER INFORMATION: SWISSPROT HIT: Q42161, EVALUATE 4.00e-51
; OTHER INFORMATION: EST HUMAN HIT: BE256272.1, EVALUATE 4.00e-50
US-09-864-761-33414
```

```
Query Match 100.0%; Score 71; DB 9; Length 97;
Best Local Similarity 100.0%; Pred. No. 8.3e-05;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY 1 IQAVLSLYASGRRTG 15
Db 15 IQAVLSLYASGRRTG 29
```

```
RESULT 2
US-10-264-049-4044
; Sequence 4044, Application US/10264049
; Publication No. US20040005579A1
; GENERAL INFORMATION:
; APPLICANT: Birse et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
; FILE REFERENCE: PA133P1
; CURRENT APPLICATION NUMBER: US/10/264,049
; CURRENT FILING DATE: 2002-10-04
; PRIOR APPLICATION NUMBER: PCT/US01/18569
; PRIOR FILING DATE: 2001-06-07
; PRIOR APPLICATION NUMBER: US 60/209,467
; PRIOR FILING DATE: 2000-06-07
; NUMBER OF SEQ ID NOS: 4360
; SOFTWARE: PatentIn Ver. 3.1
; SEQ ID NO 4044
; LENGTH: 168
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-264-049-4044
```

```
Query Match 100.0%; Score 71; DB 15; Length 168;
Best Local Similarity 100.0%; Pred. No. 0.00015;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY 1 IQAVLSLYASGRRTG 15
```

```
Db 22 IQAVLSLYASGRRTG 36
```

```
RESULT 3
US-10-437-963-169247
; Sequence 169247, Application US/10437963
; Publication No. US20040123343A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; APPLICANT: Wu, Wei
; APPLICANT: Boukharov, Andrey A.
; APPLICANT: Barbazuk, Brad
; APPLICANT: Li, Ping
; TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53221)B
; CURRENT APPLICATION NUMBER: US/10/437,963
; CURRENT FILING DATE: 2003-05-14
; NUMBER OF SEQ ID NOS: 204966
; SEQ ID NO 169247
; LENGTH: 180
; TYPE: PRT
; ORGANISM: Oryza sativa
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT4530_67685C.1.pep
US-10-437-963-169247
```

```
Query Match 100.0%; Score 71; DB 16; Length 180;
Best Local Similarity 100.0%; Pred. No. 0.00016;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY 1 IQAVLSLYASGRRTG 15
Db 71 IQAVLSLYASGRRTG 85
```

```
RESULT 4
US-10-264-049-3432
; Sequence 3432, Application US/10264049
; Publication No. US20040005579A1
; GENERAL INFORMATION:
; APPLICANT: Birse et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
; FILE REFERENCE: PA133P1
; CURRENT APPLICATION NUMBER: US/10/264,049
; CURRENT FILING DATE: 2002-10-04
; PRIOR APPLICATION NUMBER: PCT/US01/18569
; PRIOR FILING DATE: 2001-06-07
; PRIOR APPLICATION NUMBER: US 60/209,467
; PRIOR FILING DATE: 2000-06-07
; NUMBER OF SEQ ID NOS: 4360
; SOFTWARE: PatentIn Ver. 3.1
; SEQ ID NO 3432
; LENGTH: 183
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: MISC FEATURE
; LOCATION: (4)_FEATURE
; OTHER INFORMATION: Xaa equals any of the twenty naturally occurring L-amino acids
; FEATURE:
; NAME/KEY: MISC FEATURE
; LOCATION: (11)_FEATURE
; OTHER INFORMATION: Xaa equals any of the twenty naturally occurring L-amino acids
; FEATURE:
; NAME/KEY: MISC FEATURE
; LOCATION: (12)_FEATURE
; OTHER INFORMATION: Xaa equals any of the twenty naturally occurring L-amino acids
; FEATURE:
```



NAME/KEY: MISC\_FEATURE  
LOCATION: (13)  
OTHER INFORMATION: Xaa equals any of the twenty naturally occurring L-amino acids  
FEATURE:  
NAME/KEY: MISC\_FEATURE  
LOCATION: (22)  
OTHER INFORMATION: Xaa equals any of the twenty naturally occurring L-amino acids  
FEATURE:  
NAME/KEY: MISC\_FEATURE  
LOCATION: (33)  
OTHER INFORMATION: Xaa equals any of the twenty naturally occurring L-amino acids  
FEATURE:  
NAME/KEY: MISC\_FEATURE  
LOCATION: (83)  
OTHER INFORMATION: Xaa equals any of the twenty naturally occurring L-amino acids  
FEATURE:  
NAME/KEY: MISC\_FEATURE  
LOCATION: (89)  
OTHER INFORMATION: Xaa equals any of the twenty naturally occurring L-amino acids  
FEATURE:  
NAME/KEY: MISC\_FEATURE  
LOCATION: (113)  
OTHER INFORMATION: Xaa equals any of the twenty naturally occurring L-amino acids  
FEATURE:  
NAME/KEY: MISC\_FEATURE  
LOCATION: (117)  
OTHER INFORMATION: Xaa equals any of the twenty naturally occurring L-amino acids  
FEATURE:  
NAME/KEY: MISC\_FEATURE  
LOCATION: (136)  
OTHER INFORMATION: Xaa equals any of the twenty naturally occurring L-amino acids  
FEATURE:  
NAME/KEY: MISC\_FEATURE  
LOCATION: (149)  
OTHER INFORMATION: Xaa equals any of the twenty naturally occurring L-amino acids  
FEATURE:  
NAME/KEY: MISC\_FEATURE  
LOCATION: (161)  
OTHER INFORMATION: Xaa equals any of the twenty naturally occurring L-amino acids  
FEATURE:  
NAME/KEY: MISC\_FEATURE  
LOCATION: (172)  
OTHER INFORMATION: Xaa equals any of the twenty naturally occurring L-amino acids  
FEATURE:  
NAME/KEY: MISC\_FEATURE  
LOCATION: (174)  
OTHER INFORMATION: Xaa equals any of the twenty naturally occurring L-amino acids  
US-10-264-049-3432

Query Match 100.0%; Score 71; DB 15; Length 183;  
Best Local Similarity 100.0%; Pred. No. 0.00017;  
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 IQAVLSLYASGRRTTG 15  
Db 44 IQAVLSLYASGRRTTG 58  
|||||

RESULT 5  
US-10-424-599-280810  
Sequence 280810, Application US/10424599  
Publication No. US20040031072A1  
GENERAL INFORMATION:  
APPLICANT: La Rosa Thomas J  
APPLICANT: Kovalic David K  
APPLICANT: Zhou Yihua  
TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated with  
FILE REFERENCE: 38-21(53223)B  
CURRENT APPLICATION NUMBER: US/10/424,599  
CURRENT FILING DATE: 2003-04-28  
NUMBER OF SEQ ID NOS: 285684

SEQ ID NO 280810  
LENGTH: 197  
TYPE: PRT  
ORGANISM: Glycine max  
FEATURE:  
NAME/KEY: unsure  
LOCATION: (1)..  
OTHER INFORMATION: unsure at all Xaa locations  
FEATURE:  
OTHER INFORMATION: Clone ID: PAT\_MRT3847\_95594C.1.pap  
US-10-424-599-280810

Query Match 100.0%; Score 71; DB 15; Length 197;  
Best Local Similarity 100.0%; Pred. No. 0.00018;  
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 IQAVLSLYASGRRTTG 15  
Db 138 IQAVLSLYASGRRTTG 152  
|||||

RESULT 6  
US-10-264-049-4148  
Sequence 4148, Application US/10264049  
Publication No. US20040005579A1  
GENERAL INFORMATION:  
APPLICANT: Birse et al.  
TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies  
FILE REFERENCE: PA133PI  
CURRENT APPLICATION NUMBER: US/10/264,049  
CURRENT FILING DATE: 2002-10-04  
PRIOR APPLICATION NUMBER: PCT/US01/18569  
PRIOR FILING DATE: 2001-06-07  
PRIOR APPLICATION NUMBER: US 60/209,467  
PRIOR FILING DATE: 2000-06-07  
NUMBER OF SEQ ID NOS: 4360  
SOFTWARE: PatentIn Ver. 3.1  
SEQ ID NO 4148  
LENGTH: 204  
TYPE: PRT  
ORGANISM: Homo sapiens  
US-10-264-049-4148

Query Match 100.0%; Score 71; DB 15; Length 204;  
Best Local Similarity 100.0%; Pred. No. 0.00019;  
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 IQAVLSLYASGRRTTG 15  
Db 156 IQAVLSLYASGRRTTG 170  
|||||

RESULT 7  
US-10-424-599-282905  
Sequence 282905, Application US/10424599  
Publication No. US20040031072A1  
GENERAL INFORMATION:  
APPLICANT: La Rosa Thomas J  
APPLICANT: Kovalic David K  
APPLICANT: Zhou Yihua  
APPLICANT: Cao Yongwei  
TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated with  
FILE REFERENCE: 38-21(53223)B  
CURRENT APPLICATION NUMBER: US/10/424,599  
CURRENT FILING DATE: 2003-04-28  
NUMBER OF SEQ ID NOS: 285684  
SEQ ID NO 282905  
LENGTH: 208  
TYPE: PRT  
ORGANISM: Glycine max  
FEATURE:  
NAME/KEY: unsure

```
; LOCATION: (1)...(208)
; OTHER INFORMATION: unsure at all Xaa locations
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT3847_97486C.1.pep
US-10-424-599-282905

Query Match      100.0%; Score 71; DB 15; Length 208;
Best Local Similarity 100.0%; Pred. No. 0.00019;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 IQAVLSLYASGRRTG 15
    |||||
Db 138 IQAVLSLYASGRRTG 152

RESULT 8
US-10-002-631C-56
; Sequence 56, Application US/10002631C
; Publication No. US20030157486A1
; GENERAL INFORMATION:
; APPLICANT: Graff, Jonathan M.
; TITLE OF INVENTION: METHODS TO IDENTIFY SIGNAL SEQUENCES
; FILE REFERENCE: A34943 090495.0243
; CURRENT APPLICATION NUMBER: US/10/002,631C
; PRIOR FILING DATE: 2001-10-31
; PRIOR APPLICATION NUMBER: 60/300,309
; PRIOR FILING DATE: 2001-06-21
; NUMBER OF SEQ ID NOS: 324
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 56
; LENGTH: 219
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-002-631C-56

Query Match      100.0%; Score 71; DB 14; Length 219;
Best Local Similarity 100.0%; Pred. No. 0.0002;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 IQAVLSLYASGRRTG 15
    |||||
Db 163 IQAVLSLYASGRRTG 177

RESULT 9
US-10-767-701-42770
; Sequence 42770, Application US/10767701
; Publication No. US20040172694A1
; GENERAL INFORMATION:
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(53235)B
; CURRENT APPLICATION NUMBER: US/10/767,701
; CURRENT FILING DATE: 2004-01-29
; NUMBER OF SEQ ID NOS: 63128
; SEQ ID NO 42770
; LENGTH: 230
; TYPE: PRT
; ORGANISM: Sorghum bicolor
; FEATURE:
; OTHER INFORMATION: Clone ID: SORBI-28MAY03-CI7410_1.pep
US-10-767-701-42770

Query Match      100.0%; Score 71; DB 16; Length 230;
Best Local Similarity 100.0%; Pred. No. 0.00021;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 IQAVLSLYASGRRTG 15
    |||||
```

```
Db 138 IQAVLSLYASGRRTG 152

RESULT 10
US-10-424-599-206389
; Sequence 206389, Application US/10424599
; Publication No. US20040031072A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa Thomas J
; APPLICANT: Kovalic David K
; APPLICANT: Zhou Yihua
; APPLICANT: Cao Yongwei
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(53223)B
; CURRENT APPLICATION NUMBER: US/10/424,599
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 285684
; SEQ ID NO 206389
; LENGTH: 236
; TYPE: PRT
; ORGANISM: Glycine max
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (1)..(236)
; OTHER INFORMATION: unsure at all Xaa locations
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT3847_28396C.1.pep
US-10-424-599-206389

Query Match      100.0%; Score 71; DB 15; Length 236;
Best Local Similarity 100.0%; Pred. No. 0.00022;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 IQAVLSLYASGRRTG 15
    |||||
Db 6 IQAVLSLYASGRRTG 20

RESULT 11
US-10-424-599-143226
; Sequence 143226, Application US/10424599
; Publication No. US20040031072A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa Thomas J
; APPLICANT: Kovalic David K
; APPLICANT: Zhou Yihua
; APPLICANT: Cao Yongwei
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(53223)B
; CURRENT APPLICATION NUMBER: US/10/424,599
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 285684
; SEQ ID NO 143226
; LENGTH: 237
; TYPE: PRT
; ORGANISM: Glycine max
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (1)..(237)
; OTHER INFORMATION: unsure at all Xaa locations
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT3847_100347C.1.pep
US-10-424-599-143226

Query Match      100.0%; Score 71; DB 15; Length 237;
Best Local Similarity 100.0%; Pred. No. 0.00022;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 IQAVLSLYASGRRTG 15
    |||||
Db 20 IQAVLSLYASGRRTG 34
```

RESULT 12  
US-10-424-599-162684  
; Sequence 162684, Application US/10424599  
; Publication No. US20040031072A1  
; GENERAL INFORMATION:  
; APPLICANT: La Rosa Thomas J  
; APPLICANT: Kovalic David K  
; APPLICANT: Zhou Yihua  
; APPLICANT: Cao Yongwei  
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With  
; FILE OF INVENTION: Plants and Uses Thereof for Plant Improvement  
; FILE REFERENCE: 38-21(53223)B  
; CURRENT APPLICATION NUMBER: US/10/424,599  
; CURRENT FILING DATE: 2003-04-28  
; NUMBER OF SEQ ID NOS: 285684  
; SEQ ID NO 162684  
; LENGTH: 248  
; TYPE: PRT  
; ORGANISM: Glycine max  
; FEATURE:  
; OTHER INFORMATION: Clone ID: PAT\_MRT3847\_117922C.1.pep  
US-10-424-599-162684

Query Match 100.0%; Score 71; DB 15; Length 248;  
Best Local Similarity 100.0%; Pred. No. 0.00023;  
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 IQAVLSLYASGRRTTG 15  
|||  
Db 138 IQAVLSLYASGRRTTG 152  
|||

RESULT 13  
US-10-767-701-43575  
; Sequence 43575, Application US/10767701  
; Publication No. US20040172684A1  
; GENERAL INFORMATION:  
; APPLICANT: Kovalic, David K.  
; APPLICANT: Zhou, Yihua  
; APPLICANT: Cao, Yongwei  
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With  
; FILE OF INVENTION: Plants and Uses Thereof for Plant Improvement  
; FILE REFERENCE: 38-21(53535)B  
; CURRENT APPLICATION NUMBER: US/10/767,701  
; CURRENT FILING DATE: 2004-01-29  
; NUMBER OF SEQ ID NOS: 63128  
; SEQ ID NO 43575  
; LENGTH: 249  
; TYPE: PRT  
; ORGANISM: Sorghum bicolor  
; FEATURE:  
; OTHER INFORMATION: Clone ID: SORBI-28MAY03-C5949\_1.pep  
US-10-767-701-43575

Query Match 100.0%; Score 71; DB 16; Length 249;  
Best Local Similarity 100.0%; Pred. No. 0.00023;  
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 IQAVLSLYASGRRTTG 15  
|||  
Db 10 IQAVLSLYASGRRTTG 24  
|||

RESULT 14  
US-10-425-114-44494  
; Sequence 44494, Application US/10425114  
; Publication No. US20040034888A1  
; GENERAL INFORMATION:  
; APPLICANT: Liu, Jingdong  
; APPLICANT: Zhou, Yihua  
; APPLICANT: Kovalic, David K.

; APPLICANT: Screen, Steven E  
; APPLICANT: Tabaska, Jack E  
; APPLICANT: Cao, Yongwei  
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With  
; FILE OF INVENTION: Plants and Uses Thereof for Plant Improvement  
; FILE REFERENCE: 38-21(53313)B  
; CURRENT APPLICATION NUMBER: US/10/425,114  
; CURRENT FILING DATE: 2003-04-28  
; NUMBER OF SEQ ID NOS: 73128  
; SEQ ID NO 44494  
; LENGTH: 250  
; TYPE: PRT  
; ORGANISM: Glycine max  
; FEATURE:  
; OTHER INFORMATION: Clone ID: 700978095\_FLI.pep  
US-10-425-114-44494

Query Match 100.0%; Score 71; DB 15; Length 250;  
Best Local Similarity 100.0%; Pred. No. 0.00023;  
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 IQAVLSLYASGRRTTG 15  
|||  
Db 11 IQAVLSLYASGRRTTG 25  
|||

RESULT 15  
US-10-425-114-69353  
; Sequence 69353, Application US/10425114  
; Publication No. US20040034888A1  
; GENERAL INFORMATION:  
; APPLICANT: Liu, Jingdong  
; APPLICANT: Zhou, Yihua  
; APPLICANT: Kovalic, David K.  
; APPLICANT: Screen, Steven E  
; APPLICANT: Tabaska, Jack E  
; APPLICANT: Cao, Yongwei  
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With  
; FILE OF INVENTION: Plants and Uses Thereof for Plant Improvement  
; FILE REFERENCE: 38-21(53313)B  
; CURRENT APPLICATION NUMBER: US/10/425,114  
; CURRENT FILING DATE: 2003-04-28  
; NUMBER OF SEQ ID NOS: 73128  
; SEQ ID NO 69353  
; LENGTH: 254  
; TYPE: PRT  
; ORGANISM: Zea mays  
; FEATURE:  
; OTHER INFORMATION: Clone ID: UC-ZMFLB73022E07\_FLI.pep  
US-10-425-114-69353

Query Match 100.0%; Score 71; DB 15; Length 254;  
Best Local Similarity 100.0%; Pred. No. 0.00024;  
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 IQAVLSLYASGRRTTG 15  
|||  
Db 15 IQAVLSLYASGRRTTG 29  
|||

Search completed: April 8, 2005, 12:51:00  
Job time : 42.6667 secs

**THIS PAGE BLANK (USPTO)**

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: April 8, 2005, 09:58:57 ; Search time 17.5333 Seconds  
(without alignments)  
63.863 Million cell updates/sec

Title: US-09-423-351C-9  
Perfect score: 76  
Sequence: 1 DEAQSKRGILTLPYP 15

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 513545 seqs, 74649064 residues

Total number of hits satisfying chosen parameters: 513545

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : Issued Patents AA: \*  
1: /cgn2\_6/ptodata/1/iaa/5A COMB.pdp: \*  
2: /cgn2\_6/ptodata/1/iaa/5B COMB.pdp: \*  
3: /cgn2\_6/ptodata/1/iaa/6A COMB.pdp: \*  
4: /cgn2\_6/ptodata/1/iaa/6B COMB.pdp: \*  
5: /cgn2\_6/ptodata/1/iaa/ECTUS COMB.pdp: \*  
6: /cgn2\_6/ptodata/1/iaa/backfiles1.pdp: \*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	76	100.0	80	3	US-09-306-446C-4
2	76	100.0	374	3	US-08-609-236-6
3	76	100.0	374	3	US-09-306-446C-2
4	76	100.0	375	2	US-08-494-151-14
5	76	100.0	375	3	US-09-106-217-16
6	76	100.0	375	4	US-09-978-594-731
7	76	100.0	376	1	US-08-588-113-2
8	76	100.0	376	4	US-09-949-016-6100
9	76	100.0	377	3	US-09-106-217-2
10	76	100.0	377	4	US-09-919-172-33
11	76	100.0	377	4	US-09-917-254-53
12	76	100.0	386	4	US-09-949-016-7721
13	76	100.0	399	4	US-09-949-016-9424
14	76	100.0	402	4	US-09-949-016-10757
15	76	100.0	404	4	US-09-949-016-11313
16	73	96.1	371	1	US-08-261-206A-77
17	73	96.1	375	3	US-09-171-337A-7
18	73	96.1	375	3	US-09-171-337A-8
19	73	96.1	375	4	US-09-631-022-7
20	73	96.1	375	4	US-09-631-022-8
21	73	96.1	377	4	US-09-248-796A-14109
22	56	73.7	91	4	US-09-621-976-6520
23	46	60.5	376	4	US-09-538-092-1110
24	46	60.5	376	4	US-09-949-016-6656
25	46	60.5	376	4	US-09-949-016-8452
26	44	57.9	384	4	US-09-538-092-395
27	42	55.3	147	4	US-09-902-540-12005

Sequence 7725, Ap  
Sequence 1109, Ap  
Sequence 19404, A  
Sequence 11, Appl  
Sequence 4962, Ap  
Sequence 4886, Ap  
Sequence 8292, Ap  
Sequence 11500, A  
Sequence 5362, Ap  
Sequence 6555, Ap  
Sequence 7396, Ap  
Sequence 4772, Ap  
Sequence 4, Appl  
Sequence 5, Appl  
Sequence 755, App  
Sequence 67, Appl  
Sequence 6, Appl

US-09-949-016-7725  
US-09-538-092-1109  
US-09-248-796A-19404  
US-09-171-337C-11  
US-09-134-001C-4962  
US-09-107-532A-4886  
US-09-543-681A-8292  
US-09-902-540-11500  
US-09-107-532A-5362  
US-09-949-016-6555  
US-09-949-016-7396  
US-09-328-352-4772  
US-09-057-570-4  
US-09-387-693-4  
US-09-824-574-5  
US-09-732-210-755  
US-09-631-594-67  
US-08-472-172-6

ALIGNMENTS

RESULT 1  
US-09-306-446C-4  
; Sequence 4, Application US/09306446C  
; Patent No. 6372959  
; GENERAL INFORMATION:  
; APPLICANT: KIM, Dong Soo  
; APPLICANT: KIM, Chul Geun  
; APPLICANT: NAM, Yoon Kwon  
; APPLICANT: NOH, Jae Koo  
; APPLICANT: CHO, Kyou Nam  
; TITLE OF INVENTION: EXPRESSION VECTOR OF MUD LOACH GROWTH HORMONE GENE  
; FILE REFERENCE: P063440S0/BAS  
; CURRENT APPLICATION NUMBER: US/09/306.446C  
; CURRENT FILING DATE: 1999-05-06  
; PRIOR APPLICATION NUMBER: KR 98/20255  
; PRIOR FILING DATE: 1998-06-01  
; NUMBER OF SEQ ID NOS: 33  
; SOFTWARE: Patent in Ver. 2.0  
; SEQ ID NO 4  
; LENGTH: 80  
; TYPE: PRT  
; ORGANISM: Misgurnus mizolepus  
US-09-306-446C-4

Query Match 100.0%; Score 76; DB 3; Length 80;  
Best Local Similarity 100.0%; Pred. No. 4.8e-07;  
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 DEAQSKRGILTLPYP 15  
Db 15 DEAQSKRGILTLPYP 29

RESULT 2  
US-08-609-236-6  
; Sequence 6, Application US/08609236  
; Patent No. 6087398  
; GENERAL INFORMATION:  
; APPLICANT: Steven R. Goodman  
; TITLE OF INVENTION: No. 6087398el Sickle Cell Anemia Treatment  
; NUMBER OF SEQUENCES: 6  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: McGregor & Adler, LLP  
; STREET: 8011 Candle Lane  
; CITY: Houston  
; STATE: Texas  
; COUNTRY: USA  
; ZIP: 77071  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk

COMPUTER: Apple Macintosh  
OPERATING SYSTEM: Macintosh  
SOFTWARE: Microsoft Word for Macintosh  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/609,236  
FILING DATE: March 1, 1996  
CLASSIFICATION: 514  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 60/002,288  
FILING DATE: August 14, 1995  
CLASSIFICATION: 514  
ATTORNEY/AGENT INFORMATION:  
NAME: Benjamin Aaron Adler, Ph.D.  
REGISTRATION NUMBER: 35,423  
REFERENCE/DOCKET NUMBER: D5807  
TELEPHONE: 713-777-2321  
TELEFAX: 713-777-6908  
INFORMATION FOR SEQ ID NO: 6:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 374  
TYPE: Amino acid  
STRANDEDNESS:  
TOPOLOGY: Linear  
MOLECULE TYPE: Protein  
DESCRIPTION: Protein  
HYPOTHETICAL: No  
ANTI-SENSE: No  
FRAGMENT TYPE:  
ORIGINAL SOURCE:  
STRAIN:  
INDIVIDUAL ISOLATE:  
DEVELOPMENTAL STAGE:  
TISSUE TYPE:  
CELL TYPE:  
CELL LINE:  
US-08-609-236-6

Query Match 100.0%; Score 76; DB 3; Length 374;  
Best Local Similarity 100.0%; Pred. No. 2.6e-06;  
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 DEAQSKRGILTLYKYP 15  
Db 55 DEAQSKRGILTLYKYP 69

RESULT 3  
US-09-306-446C-2  
Sequence 2, Application US/09306446C  
Patent No. 6372959  
GENERAL INFORMATION:  
APPLICANT: KIM, Dong Soo  
APPLICANT: KIM, Chul Geun  
APPLICANT: NAM, Yoon Kwon  
APPLICANT: NOH, Jae Koo  
APPLICANT: CHO, Kyou Nam  
TITLE OF INVENTION: EXPRESSION VECTOR OF MUD LOACH GROWTH HORMONE GENE  
FILE REFERENCE: P06344USO/BAS  
CURRENT APPLICATION NUMBER: US/09/306,446C  
CURRENT FILING DATE: 1999-05-06  
PRIOR APPLICATION NUMBER: KR 98/20255  
PRIOR FILING DATE: 1998-06-01  
NUMBER OF SEQ ID NOS: 33  
SOFTWARE: Patent In Ver. 2.0  
SEQ ID NO 2  
LENGTH: 374  
TYPE: PRT  
ORGANISM: Misgurnus mizolepus  
US-09-306-446C-2

Query Match 100.0%; Score 76; DB 3; Length 374;  
Best Local Similarity 100.0%; Pred. No. 2.6e-06;

Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
Qy 1 DEAQSKRGILTLYKYP 15  
Db 56 DEAQSKRGILTLYKYP 70

RESULT 4  
US-08-494-151-14  
Sequence 14, Application US/08494151  
Patent No. 5840528  
GENERAL INFORMATION:  
APPLICANT: Van Ooyen, Albert Johannes Joseph  
TITLE OF INVENTION: Transformation of Phaffia rhodozyma  
NUMBER OF SEQUENCES: 14  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Morrison & Foerster  
STREET: 2000 Pennsylvania Ave. N.W., Suite 5500  
CITY: Washington, D.C.  
COUNTRY: USA  
ZIP: 20006-1812  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent In Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/494,151  
FILING DATE: 23-JUN-1995  
CLASSIFICATION: 435  
ATTORNEY/AGENT INFORMATION:  
NAME: Murashige, Kate H.  
REGISTRATION NUMBER: 29,959  
REFERENCE/DOCKET NUMBER: 24615-20039.01  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (202) 887-1500  
TELEFAX: (202) 887-0763  
TELEX: 90-4030  
INFORMATION FOR SEQ ID NO: 14:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 375 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-494-151-14

Query Match 100.0%; Score 76; DB 2; Length 375;  
Best Local Similarity 100.0%; Pred. No. 2.6e-06;  
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 DEAQSKRGILTLYKYP 15  
Db 56 DEAQSKRGILTLYKYP 70

RESULT 5  
US-09-106-217-16  
Sequence 16, Application US/09106217  
Patent No. 6063576  
GENERAL INFORMATION:  
APPLICANT: Keating, Mark T.  
APPLICANT: Olson, Timothy M.  
TITLE OF INVENTION: Actin Mutations in Dilated  
TITLE OF INVENTION: Cardiomyopathy, a Heritable Form of Heart Failure  
NUMBER OF SEQUENCES: 18  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Rothwell, Figg, Ernst & Kurz, P.C.  
STREET: 555 Thirteenth Street, N.W., Suite 701 East  
CITY: Washington  
STATE: DC  
COUNTRY: U.S.A.  
ZIP: 20004

COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
FILING DATE: US/09/106,217  
CLASSIFICATION:  
ATTORNEY/AGENT INFORMATION:  
NAME: Saxe, Stephen A.  
REGISTRATION NUMBER: 38,609  
REFERENCE/DOCKET NUMBER: 2323-125  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 202-783-6040  
TELEFAX: 202-783-6031  
INFORMATION FOR SEQ ID NO: 16:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 375 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-09-106-217-16

Query Match 100.0%; Score 76; DB 3; Length 375;  
Best Local Similarity 100.0%; Pred. No. 2.6e-06;  
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 DEAQSKRGILTLYKYP 15  
|||||  
Db 56 DEAQSKRGILTLYKYP 70

## RESULT 6

US-09-976-594-731  
Sequence 731, Application US/09976594  
Patent No. 6673549

GENERAL INFORMATION:  
APPLICANT: Furness, Michael  
APPLICANT: Buchbinder, Jenny  
TITLE OF INVENTION: GENES EXPRESSED IN C3A LIVER CELL CULTURES TREATED WITH STEROIDS  
FILE REFERENCE: PA-0041 US  
CURRENT APPLICATION NUMBER: US/09/976,594  
CURRENT FILING DATE: 2001-10-12  
PRIOR APPLICATION NUMBER: 60/240,409  
PRIOR FILING DATE: 2000-10-12  
NUMBER OF SEQ ID NOS: 1143  
SOFTWARE: PERL Program  
SEQ ID NO 731  
LENGTH: 375  
TYPE: PRT  
ORGANISM: Homo sapiens  
FEATURE:  
NAME/KEY: misc feature  
OTHER INFORMATION: Incyte ID No. 6673549 1837317CD1  
US-09-976-594-731

Query Match 100.0%; Score 76; DB 4; Length 375;  
Best Local Similarity 100.0%; Pred. No. 2.6e-06;  
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 DEAQSKRGILTLYKYP 15  
|||||  
Db 56 DEAQSKRGILTLYKYP 70

## RESULT 7

US-08-588-113-2  
Sequence 2, Application US/08588113  
Patent No. 5710003

GENERAL INFORMATION:  
APPLICANT: McHugh, Kirk M.  
TITLE OF INVENTION: DIAGNOSTIC TEST FOR DETERMINING

TITLE OF INVENTION: MALIGNANCY OF SMOOTH MUSCLE TUMORS  
NUMBER OF SEQUENCES: 4  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Woodcock Washburn Kurtz Mackiewicz &  
ADDRESSEE: No. 5710003ris  
STREET: One Liberty Place - 46th Floor  
CITY: Philadelphia  
STATE: PA  
COUNTRY: USA  
ZIP: 19103  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/588,113  
FILING DATE:  
CLASSIFICATION: 436  
ATTORNEY/AGENT INFORMATION:  
NAME: Ralph, Rebecca L.  
REGISTRATION NUMBER: 35,152  
REFERENCE/DOCKET NUMBER: TJU-1652  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 215-568-3100  
TELEFAX: 215-568-3439  
INFORMATION FOR SEQ ID NO: 2:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 376 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-588-113-2

Query Match 100.0%; Score 76; DB 1; Length 376;  
Best Local Similarity 100.0%; Pred. No. 2.6e-06;  
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 DEAQSKRGILTLYKYP 15  
|||||  
Db 57 DEAQSKRGILTLYKYP 71

## RESULT 8

US-09-949-016-6100  
Sequence 6100, Application US/09949016  
Patent No. 6812339

GENERAL INFORMATION:  
APPLICANT: VENTER, J. Craig et al.  
TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF  
FILE REFERENCE: CL001307  
CURRENT APPLICATION NUMBER: US/09/949,016  
CURRENT FILING DATE: 2000-04-14  
PRIOR APPLICATION NUMBER: 60/241,755  
PRIOR FILING DATE: 2000-10-20  
PRIOR APPLICATION NUMBER: 60/237,768  
PRIOR FILING DATE: 2000-10-03  
PRIOR APPLICATION NUMBER: 60/231,498  
PRIOR FILING DATE: 2000-09-08  
NUMBER OF SEQ ID NOS: 207012  
SOFTWARE: FastSeq for Windows Version 4.0  
SEQ ID NO 6100  
LENGTH: 376  
TYPE: PRT  
ORGANISM: Human  
US-09-949-016-6100

Query Match 100.0%; Score 76; DB 4; Length 376;  
Best Local Similarity 100.0%; Pred. No. 2.6e-06;  
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 DEAQSKRGILTLYKYP 15

```
Db 57 DEAQSKRGILTLYKYP 71
|||||
RESULT 9
US-09-106-217-2
; Sequence 2, Application US/09106217
; Patent No. 6063576
; GENERAL INFORMATION:
; APPLICANT: Keating, Mark T.
; TITLE OF INVENTION: Actin Mutations in Dilated
; TITLE OF INVENTION: Cardiomyopathy, a Heritable Form of Heart Failure
; NUMBER OF SEQUENCES: 18
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Rothwell, Figg, Ernst & Kurz, P.C.
; STREET: 555 Thirteenth Street, N.W., Suite 701 East
; STREET: Tower
; CITY: Washington
; STATE: DC
; COUNTRY: U.S.A.
; ZIP: 20004
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/106,217
; FILING DATE:
; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
; NAME: Saxe, Stephen A.
; REGISTRATION NUMBER: 38,609
; REFERENCE/DOCKET NUMBER: 2323-125
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 202-783-6040
; TELEFAX: 202-783-6031
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 377 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-09-106-217-2

Query Match 100.0%; Score 76; DB 3; Length 377;
Best Local Similarity 100.0%; Pred. No. 2.6e-06;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 DEAQSKRGILTLYKYP 15
|||||
Db 58 DEAQSKRGILTLYKYP 72
|||||
RESULT 10
US-09-919-172-33
; Sequence 33, Application US/09919172
; Patent No. 6673545
; GENERAL INFORMATION:
; APPLICANT: Faris, Mary
; APPLICANT: Turner, Christopher M.
; TITLE OF INVENTION: PROSTATE CANCER MARKERS
; FILE REFERENCE: PA-0036 US
; CURRENT APPLICATION NUMBER: US/09/919,172
; CURRENT FILING DATE: 2001-07-30
; PRIOR APPLICATION NUMBER: 60/222,469
; PRIOR FILING DATE: 2000-07-28
; NUMBER OF SEQ ID NOS: 102
; SOFTWARE: PERL Program
; SEQ ID NO 33
; LENGTH: 377
; TYPE: PRT
; ORGANISM: Human
; US-09-919-172-33

Query Match 100.0%; Score 76; DB 4; Length 386;
Best Local Similarity 100.0%; Pred. No. 2.7e-06;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 DEAQSKRGILTLYKYP 15
|||||
Db 58 DEAQSKRGILTLYKYP 72
|||||
RESULT 11
US-09-917-254-53
; Sequence 53, Application US/09917254
; Patent No. 6703204
; GENERAL INFORMATION:
; APPLICANT: Mutter, George
; APPLICANT: Baak, Jan
; TITLE OF INVENTION: Prognostic Classification of Breast Cancer
; FILE REFERENCE: B0801/7224(JRV)
; CURRENT APPLICATION NUMBER: US/09/917,254
; CURRENT FILING DATE: 2001-07-27
; PRIOR APPLICATION NUMBER: US 60/222,093
; PRIOR FILING DATE: 2000-07-28
; NUMBER OF SEQ ID NOS: 102
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 53
; LENGTH: 377
; TYPE: PRT
; ORGANISM: Homo Sapiens
; US-09-917-254-53

Query Match 100.0%; Score 76; DB 4; Length 377;
Best Local Similarity 100.0%; Pred. No. 2.6e-06;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 DEAQSKRGILTLYKYP 15
|||||
Db 58 DEAQSKRGILTLYKYP 72
|||||
RESULT 12
US-09-949-016-7721
; Sequence 7721, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; TITLE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 7721
; LENGTH: 386
; TYPE: PRT
; ORGANISM: Human
; US-09-949-016-7721

Query Match 100.0%; Score 76; DB 4; Length 386;
Best Local Similarity 100.0%; Pred. No. 2.7e-06;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```



Qy 1 DEAQSKRGILTLYP 15  
| | | | | | | | | |  
Db 67 DEAQSKRGILTLYP 81

RESULT 13  
US-09-949-016-9424  
; Sequence 9424, Application US/09949016  
; Patent No. 6812339  
; GENERAL INFORMATION:  
; APPLICANT: VENTER, J. Craig et al.  
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED  
; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF  
; FILE REFERENCE: CL001307  
; CURRENT APPLICATION NUMBER: US/09/949,016  
; CURRENT FILING DATE: 2000-04-14  
; PRIOR APPLICATION NUMBER: 60/241,755  
; PRIOR FILING DATE: 2000-10-20  
; PRIOR APPLICATION NUMBER: 60/237,768  
; PRIOR FILING DATE: 2000-10-03  
; PRIOR APPLICATION NUMBER: 60/231,498  
; PRIOR FILING DATE: 2000-09-08  
; NUMBER OF SEQ ID NOS: 207012  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 9424  
; LENGTH: 399  
; TYPE: PRT  
; ORGANISM: Human  
US-09-949-016-9424

Query Match 100.0%; Score 76; DB 4; Length 399;  
Best Local Similarity 100.0%; Pred. No. 2.8e-06;  
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 DEAQSKRGILTLYP 15  
| | | | | | | | | |  
Db 80 DEAQSKRGILTLYP 94

RESULT 14  
US-09-949-016-10757  
; Sequence 10757, Application US/09949016  
; Patent No. 6812339  
; GENERAL INFORMATION:  
; APPLICANT: VENTER, J. Craig et al.  
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED  
; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF  
; FILE REFERENCE: CL001307  
; CURRENT APPLICATION NUMBER: US/09/949,016  
; CURRENT FILING DATE: 2000-04-14  
; PRIOR APPLICATION NUMBER: 60/241,755  
; PRIOR FILING DATE: 2000-10-20  
; PRIOR APPLICATION NUMBER: 60/237,768  
; PRIOR FILING DATE: 2000-10-03  
; PRIOR APPLICATION NUMBER: 60/231,498  
; PRIOR FILING DATE: 2000-09-08  
; NUMBER OF SEQ ID NOS: 207012  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 10757  
; LENGTH: 402  
; TYPE: PRT  
; ORGANISM: Human  
US-09-949-016-10757

Query Match 100.0%; Score 76; DB 4; Length 402;  
Best Local Similarity 100.0%; Pred. No. 2.8e-06;  
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 DEAQSKRGILTLYP 15  
| | | | | | | | | |  
Db 83 DEAQSKRGILTLYP 97

RESULT 15  
US-09-949-016-11313  
; Sequence 11313, Application US/09949016  
; Patent No. 6812339  
; GENERAL INFORMATION:  
; APPLICANT: VENTER, J. Craig et al.  
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED  
; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF  
; FILE REFERENCE: CL001307  
; CURRENT APPLICATION NUMBER: US/09/949,016  
; CURRENT FILING DATE: 2000-04-14  
; PRIOR APPLICATION NUMBER: 60/241,755  
; PRIOR FILING DATE: 2000-10-20  
; PRIOR APPLICATION NUMBER: 60/237,768  
; PRIOR FILING DATE: 2000-10-03  
; PRIOR APPLICATION NUMBER: 60/231,498  
; PRIOR FILING DATE: 2000-09-08  
; NUMBER OF SEQ ID NOS: 207012  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 11313  
; LENGTH: 404  
; TYPE: PRT  
; ORGANISM: Human  
US-09-949-016-11313

Query Match 100.0%; Score 76; DB 4; Length 404;  
Best Local Similarity 100.0%; Pred. No. 2.8e-06;  
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 DEAQSKRGILTLYP 15  
| | | | | | | | | |  
Db 85 DEAQSKRGILTLYP 99

Search completed: April 8, 2005, 12:07:37  
Job time : 17.5333 secs

**THIS PAGE BLANK (USPTO)**

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

## OM protein - protein search, using sw model

Run on: April 8, 2005, 10:53:18 ; Search time 42.6667 Seconds  
(without alignments)  
116.718 Million cell updates/sec

Title: US-09-423-351C-9

Perfect score: 76

Sequence: 1 DEAQSKRGILTLYKYP 15

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1418010 seqs, 331997259 residues

Total number of hits satisfying chosen parameters: 1418010

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Listing first 45 summaries

## Database :

Published Applications AA:\*

- 1: /cgn2\_6/ptodata/1/pubpaa/US07\_PUBCOMB.pep.\*
- 2: /cgn2\_6/ptodata/1/pubpaa/PCT\_NEW\_PUB.pep.\*
- 3: /cgn2\_6/ptodata/1/pubpaa/US05\_NEW\_PUB.pep.\*
- 4: /cgn2\_6/ptodata/1/pubpaa/US06\_PUBCOMB.pep.\*
- 5: /cgn2\_6/ptodata/1/pubpaa/US07\_NEW\_PUB.pep.\*
- 6: /cgn2\_6/ptodata/1/pubpaa/PCTUS\_PUBCOMB.pep.\*
- 7: /cgn2\_6/ptodata/1/pubpaa/US08\_NEW\_PUB.pep.\*
- 8: /cgn2\_6/ptodata/1/pubpaa/US08\_PUBCOMB.pep.\*
- 9: /cgn2\_6/ptodata/1/pubpaa/US09A\_PUBCOMB.pep.\*
- 10: /cgn2\_6/ptodata/1/pubpaa/US09B\_PUBCOMB.pep.\*
- 11: /cgn2\_6/ptodata/1/pubpaa/US09C\_PUBCOMB.pep.\*
- 12: /cgn2\_6/ptodata/1/pubpaa/US09C\_NEW\_PUB.pep.\*
- 13: /cgn2\_6/ptodata/1/pubpaa/US10A\_PUBCOMB.pep.\*
- 14: /cgn2\_6/ptodata/1/pubpaa/US10B\_PUBCOMB.pep.\*
- 15: /cgn2\_6/ptodata/1/pubpaa/US10C\_PUBCOMB.pep.\*
- 16: /cgn2\_6/ptodata/1/pubpaa/US10D\_PUBCOMB.pep.\*
- 17: /cgn2\_6/ptodata/1/pubpaa/US10\_NEW\_PUB.pep.\*
- 18: /cgn2\_6/ptodata/1/pubpaa/US11\_NEW\_PUB.pep.\*
- 19: /cgn2\_6/ptodata/1/pubpaa/US60\_NEW\_PUB.pep.\*
- 20: /cgn2\_6/ptodata/1/pubpaa/US60\_PUBCOMB.pep.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	76	100.0	105	US-10-264-049-3601	Sequence 3601, Ap
2	76	100.0	108	US-09-864-761-37942	Sequence 37942, A
3	76	100.0	135	US-10-767-701-47318	Sequence 47318, A
4	76	100.0	136	US-10-424-599-280811	Sequence 280811,
5	76	100.0	142	US-10-424-599-184471	Sequence 184471,
6	76	100.0	197	US-10-424-599-280810	Sequence 280810,
7	76	100.0	201	US-10-264-049-4308	Sequence 4308, Ap
8	76	100.0	204	US-10-264-049-4148	Sequence 4148, Ap
9	76	100.0	208	US-10-424-599-282905	Sequence 282905,
10	76	100.0	219	US-10-002-631C-56	Sequence 56, Appl
11	76	100.0	219	US-10-029-386-32056	Sequence 32056, A
12	76	100.0	230	US-10-767-701-42770	Sequence 42770, A
13	76	100.0	248	US-10-424-599-162684	Sequence 162684,

14	76	100.0	331	US-10-425-114-59172	Sequence 59172, A
15	76	100.0	342	US-10-108-260A-4008	Sequence 4008, Ap
16	76	100.0	347	US-10-424-599-283333	Sequence 283333,
17	76	100.0	371	US-10-322-281-68	Sequence 68, Appl
18	76	100.0	375	US-10-205-194-93	Sequence 93, Appl
19	76	100.0	375	US-10-316-253-88	Sequence 88, Appl
20	76	100.0	375	US-10-369-493-2436	Sequence 2436, Ap
21	76	100.0	375	US-10-369-493-5927	Sequence 5927, Ap
22	76	100.0	375	US-10-205-331-94	Sequence 94, Appl
23	76	100.0	375	US-10-260-708-82	Sequence 82, Appl
24	76	100.0	376	US-10-341-434-101	Sequence 101, App
25	76	100.0	376	US-10-437-963-202420	Sequence 202420, A
26	76	100.0	376	US-10-767-701-45677	Sequence 45677, A
27	76	100.0	377	US-09-919-172-33	Sequence 33, Appl
28	76	100.0	377	US-10-338-777-52	Sequence 52, Appl
29	76	100.0	377	US-10-236-031B-64	Sequence 64, Appl
30	76	100.0	377	US-10-424-599-162009	Sequence 162009,
31	76	100.0	377	US-10-424-599-162685	Sequence 162685,
32	76	100.0	377	US-10-424-599-162686	Sequence 162686,
33	76	100.0	377	US-10-424-599-222816	Sequence 222816,
34	76	100.0	377	US-10-424-599-223492	Sequence 223492,
35	76	100.0	377	US-10-424-599-283336	Sequence 283336,
36	76	100.0	377	US-10-425-114-66124	Sequence 66124, A
37	76	100.0	377	US-10-333-680-4	Sequence 4, Appli
38	76	100.0	377	US-10-437-963-121952	Sequence 121952,
39	76	100.0	377	US-10-437-963-148877	Sequence 148877,
40	76	100.0	377	US-10-437-963-198295	Sequence 198295,
41	76	100.0	377	US-10-767-701-47239	Sequence 47239, A
42	76	100.0	378	US-10-425-114-42317	Sequence 42317, A
43	76	100.0	378	US-10-425-114-52458	Sequence 52458, A
44	76	100.0	378	US-10-425-114-59265	Sequence 59265, A
45	76	100.0	378	US-10-425-114-59967	Sequence 59967, A

## ALIGNMENTS

## RESULT 1

US-10-264-049-3601  
; Sequence 3601, Application US/10264049  
; Publication No. US20040005579A1  
; GENERAL INFORMATION:  
; APPLICANT: Birse et al.  
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies  
; FILE REFERENCE: PAL33P1  
; CURRENT APPLICATION NUMBER: US/10/264,049  
; CURRENT FILING DATE: 2002-10-04  
; PRIOR APPLICATION NUMBER: PCT/US01/18569  
; PRIOR FILING DATE: 2001-06-07  
; PRIOR APPLICATION NUMBER: US 60/209,467  
; PRIOR FILING DATE: 2000-06-07  
; NUMBER OF SEQ ID NOS: 4360  
; SOFTWARE: Patent In Ver. 3.1  
; SEQ ID NO 3601  
; LENGTH: 105  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
; FEATURE:  
; NAME/KEY: MISC\_FEATURE  
; LOCATION: (103)  
; OTHER INFORMATION: Xaa equals any of the twenty naturally occurring L-amino acids  
US-10-264-049-3601

Query Match 100.0%; Score 76; DB 15; Length 105;  
Best Local Similarity 100.0%; Pred. No. 5e-06;  
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 DEAQSKRGILTLYKYP 15

Db 57 DEAQSKRGILTLYKYP 71

## RESULT 2

US-09-864-761-37942  
; Sequence 37942, Application US/09864761  
; Patent No. US20020048763A1  
; GENERAL INFORMATION:  
; APPLICANT: Penn, Sharron G.  
; APPLICANT: Rank, David R.  
; APPLICANT: Hanzel, David K.  
; APPLICANT: Chen, Wensheng  
; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR  
; TITLE OF INVENTION: GENE EXPRESSION ANALYSIS BY MICROARRAY  
; FILE REFERENCE: Aecm1ca-X-1  
; CURRENT APPLICATION NUMBER: US/09/864,761  
; CURRENT FILING DATE: 2001-05-23  
; PRIOR APPLICATION NUMBER: US 60/180,312  
; PRIOR FILING DATE: 2000-02-04  
; PRIOR APPLICATION NUMBER: US 60/207,456  
; PRIOR FILING DATE: 2000-05-26  
; PRIOR APPLICATION NUMBER: US 09/632,366  
; PRIOR FILING DATE: 2000-08-03  
; PRIOR APPLICATION NUMBER: GB 24263.6  
; PRIOR FILING DATE: 2000-10-04  
; PRIOR APPLICATION NUMBER: US 60/236,359  
; PRIOR FILING DATE: 2000-09-27  
; PRIOR APPLICATION NUMBER: PCT/US01/00666  
; PRIOR FILING DATE: 2001-01-30  
; PRIOR APPLICATION NUMBER: PCT/US01/00667  
; PRIOR FILING DATE: 2001-01-30  
; PRIOR APPLICATION NUMBER: PCT/US01/00664  
; PRIOR FILING DATE: 2001-01-30  
; PRIOR APPLICATION NUMBER: PCT/US01/00669  
; PRIOR FILING DATE: 2001-01-30  
; PRIOR APPLICATION NUMBER: PCT/US01/00665  
; PRIOR FILING DATE: 2001-01-30  
; PRIOR APPLICATION NUMBER: PCT/US01/00668  
; PRIOR FILING DATE: 2001-01-30  
; PRIOR APPLICATION NUMBER: PCT/US01/00663  
; PRIOR FILING DATE: 2001-01-30  
; PRIOR APPLICATION NUMBER: PCT/US01/00662  
; PRIOR FILING DATE: 2001-01-30  
; PRIOR APPLICATION NUMBER: PCT/US01/00661  
; PRIOR FILING DATE: 2001-01-30  
; PRIOR APPLICATION NUMBER: PCT/US01/00670  
; PRIOR FILING DATE: 2001-01-30  
; PRIOR APPLICATION NUMBER: US 60/234,687  
; PRIOR FILING DATE: 2000-09-21  
; PRIOR APPLICATION NUMBER: US 09/608,408  
; PRIOR FILING DATE: 2000-06-30  
; PRIOR APPLICATION NUMBER: US 09/774,203  
; PRIOR FILING DATE: 2001-01-29  
; NUMBER OF SEQ ID NOS: 49117  
; SOFTWARE: Anomax Sequence Listing Engine vers. 1.1  
; SEQ ID NO 37942  
; LENGTH: 108  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
; FEATURE:  
; OTHER INFORMATION: MAP TO AC012271.1  
; OTHER INFORMATION: EXPRESSED IN HEL100, SIGNAL = 14  
; OTHER INFORMATION: EXPRESSED IN HEL100, SIGNAL = 17  
; OTHER INFORMATION: EXPRESSED IN HEART, SIGNAL = 62  
; OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 13  
; OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 28  
; OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 9.9  
; OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 3.1  
; OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 12  
; OTHER INFORMATION: EXPRESSED IN BT474, SIGNAL = 27  
; OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 14  
; OTHER INFORMATION: SWISSPROT HIT: O15998, EVALUE 9.00e-60  
; OTHER INFORMATION: EST\_HUMAN HIT: BF127509.1, EVALUE 2.00e-58

US-09-864-761-37942  
Query Match 100.0%; Score 76; DB 9; Length 108;  
Best Local Similarity 100.0%; Pred. No. 5.2e-06;

Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
Qy 1 DEAQSKRGILTLYP 15  
Db 15 DEAQSKRGILTLYP 29  
|||||

## RESULT 3

US-10-767-701-47318  
; Sequence 47318, Application US/10767701  
; Publication No. US20040172684A1  
; GENERAL INFORMATION:  
; APPLICANT: Kovalic, David K.  
; APPLICANT: Zhou, Yihua  
; APPLICANT: Cao, Yongwei  
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated with  
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement  
; FILE REFERENCE: 38-21(53535)B  
; CURRENT APPLICATION NUMBER: US/10/767,701  
; CURRENT FILING DATE: 2004-01-29  
; NUMBER OF SEQ ID NOS: 63128  
; SEQ ID NO 47318  
; LENGTH: 135  
; TYPE: PRT  
; ORGANISM: Sorghum bicolor  
; FEATURE:  
; NAME/KEY: unsure  
; LOCATION: (1)..(135)  
; OTHER INFORMATION: unsure at all xaa locations  
; FEATURE:  
; OTHER INFORMATION: Clone ID: SORBI-28MAY03-C105\_2.pep  
US-10-767-701-47318

Query Match 100.0%; Score 76; DB 16; Length 135;  
Best Local Similarity 100.0%; Pred. No. 6.6e-06;  
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 DEAQSKRGILTLYP 15  
Db 58 DEAQSKRGILTLYP 72  
|||||

## RESULT 4

US-10-424-599-280811  
; Sequence 280811, Application US/10424599  
; Publication No. US20040031072A1  
; GENERAL INFORMATION:  
; APPLICANT: La Rosa Thomas J  
; APPLICANT: Kovalic David K  
; APPLICANT: Zhou Yihua  
; APPLICANT: Cao Yongwei  
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated with  
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement  
; FILE REFERENCE: 38-21(53223)B  
; CURRENT APPLICATION NUMBER: US/10/424,599  
; CURRENT FILING DATE: 2003-04-28  
; NUMBER OF SEQ ID NOS: 285684  
; SEQ ID NO 280811  
; LENGTH: 136  
; TYPE: PRT  
; ORGANISM: Glycine max  
; FEATURE:  
; OTHER INFORMATION: Clone ID: PAT\_MRT3847\_95595C.1.pep  
US-10-424-599-280811

Query Match 100.0%; Score 76; DB 15; Length 136;  
Best Local Similarity 100.0%; Pred. No. 6.7e-06;  
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 DEAQSKRGILTLYP 15  
Db 58 DEAQSKRGILTLYP 72  
|||||

RESULT 5  
US-10-424-599-184471  
; Sequence 184471, Application US/10424599  
; Publication No. US20040031072A1  
; GENERAL INFORMATION:  
; APPLICANT: La Rosa Thomas J  
; APPLICANT: Kovalic David K  
; APPLICANT: Zhou Yihua  
; APPLICANT: Cao Yongwei  
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated with  
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement  
; FILE REFERENCE: 38-21(53223)B  
; CURRENT APPLICATION NUMBER: US/10/424,599  
; CURRENT FILING DATE: 2003-04-28  
; NUMBER OF SEQ ID NOS: 285684  
; SEQ ID NO 184471  
; LENGTH: 142  
; TYPE: PRT  
; ORGANISM: Glycine max  
; FEATURE:  
; OTHER INFORMATION: Clone ID: PAT\_MRT3847\_137594C.1.pep  
US-10-424-599-184471

Query Match 100.0%; Score 76; DB 15; Length 142;  
Best Local Similarity 100.0%; Pred. No. 7e-06;  
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 DEAQSKRGILTLYKYP 15  
|||||  
Db 65 DEAQSKRGILTLYKYP 79

RESULT 6  
US-10-424-599-280810  
; Sequence 280810, Application US/10424599  
; Publication No. US20040031072A1  
; GENERAL INFORMATION:  
; APPLICANT: La Rosa Thomas J  
; APPLICANT: Kovalic David K  
; APPLICANT: Zhou Yihua  
; APPLICANT: Cao Yongwei  
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated with  
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement  
; FILE REFERENCE: 38-21(53223)B  
; CURRENT APPLICATION NUMBER: US/10/424,599  
; CURRENT FILING DATE: 2003-04-28  
; NUMBER OF SEQ ID NOS: 285684  
; SEQ ID NO 280810  
; LENGTH: 197  
; TYPE: PRT  
; ORGANISM: Glycine max  
; FEATURE:  
; NAME/KEY: unsure  
; LOCATION: (1)..(197)  
; OTHER INFORMATION: unsure at all Xaa locations  
; FEATURE:  
; OTHER INFORMATION: Clone ID: PAT\_MRT3847\_95594C.1.pep  
US-10-424-599-280810

Query Match 100.0%; Score 76; DB 15; Length 197;  
Best Local Similarity 100.0%; Pred. No. 1e-05;  
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 DEAQSKRGILTLYKYP 15  
|||||  
Db 58 DEAQSKRGILTLYKYP 72

RESULT 7  
US-10-264-049-4308  
; Sequence 4308, Application US/10264049  
; Publication No. US20040005579A1

; GENERAL INFORMATION:  
; APPLICANT: Birse et al.  
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies  
; FILE REFERENCE: PA133P1  
; CURRENT APPLICATION NUMBER: US/10/264,049  
; CURRENT FILING DATE: 2002-10-04  
; PRIOR APPLICATION NUMBER: PCT/US01/18569  
; PRIOR FILING DATE: 2001-06-07  
; PRIOR APPLICATION NUMBER: US 60/209,467  
; PRIOR FILING DATE: 2000-06-07  
; NUMBER OF SEQ ID NOS: 4360  
; SOFTWARE: PatentIn Ver. 3.1  
; SEQ ID NO 4308  
; LENGTH: 201  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
; FEATURE:  
; NAME/KEY: MISC\_FEATURE  
; LOCATION: (135)  
; OTHER INFORMATION: Xaa equals any of the twenty naturally occurring L-amino acids  
; FEATURE:  
; NAME/KEY: MISC\_FEATURE  
; LOCATION: (136)  
; OTHER INFORMATION: Xaa equals any of the twenty naturally occurring L-amino acids  
; FEATURE:  
; NAME/KEY: MISC\_FEATURE  
; LOCATION: (140)  
; OTHER INFORMATION: Xaa equals any of the twenty naturally occurring L-amino acids  
; FEATURE:  
; NAME/KEY: MISC\_FEATURE  
; LOCATION: (143)  
; OTHER INFORMATION: Xaa equals any of the twenty naturally occurring L-amino acids  
; FEATURE:  
; NAME/KEY: MISC\_FEATURE  
; LOCATION: (144)  
; OTHER INFORMATION: Xaa equals any of the twenty naturally occurring L-amino acids  
; FEATURE:  
; NAME/KEY: MISC\_FEATURE  
; LOCATION: (145)  
; OTHER INFORMATION: Xaa equals any of the twenty naturally occurring L-amino acids  
; FEATURE:  
; NAME/KEY: MISC\_FEATURE  
; LOCATION: (151)  
; OTHER INFORMATION: Xaa equals any of the twenty naturally occurring L-amino acids  
; FEATURE:  
; NAME/KEY: MISC\_FEATURE  
; LOCATION: (159)  
; OTHER INFORMATION: Xaa equals any of the twenty naturally occurring L-amino acids  
; FEATURE:  
; NAME/KEY: MISC\_FEATURE  
; LOCATION: (172)  
; OTHER INFORMATION: Xaa equals any of the twenty naturally occurring L-amino acids  
; FEATURE:  
; NAME/KEY: MISC\_FEATURE  
; LOCATION: (174)  
; OTHER INFORMATION: Xaa equals any of the twenty naturally occurring L-amino acids  
; FEATURE:  
; NAME/KEY: MISC\_FEATURE  
; LOCATION: (180)  
; OTHER INFORMATION: Xaa equals any of the twenty naturally occurring L-amino acids  
US-10-264-049-4308

Query Match 100.0%; Score 76; DB 15; Length 201;  
Best Local Similarity 100.0%; Pred. No. 1e-05;  
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 DEAQSKRGILTLYKYP 15  
|||||  
Db 73 DEAQSKRGILTLYKYP 87

RESULT 8  
US-10-264-049-4148

; Sequence 4148, Application US/10264049  
; Publication No. US20040005579A1  
; GENERAL INFORMATION:  
; APPLICANT: Birse et al.  
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies  
; FILE REFERENCE: PA133P1  
; CURRENT APPLICATION NUMBER: US/10/264,049  
; CURRENT FILING DATE: 2002-10-04  
; PRIOR APPLICATION NUMBER: PCT/US01/18569  
; PRIOR FILING DATE: 2001-06-07  
; PRIOR APPLICATION NUMBER: US 60/209,467  
; PRIOR FILING DATE: 2000-06-07  
; NUMBER OF SEQ ID NOS: 4360  
; SOFTWARE: PatentIn Ver. 3.1  
; SEQ ID NO 4148  
; LENGTH: 204  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-10-264-049-4148

Query Match 100.0%; Score 76; DB 15; Length 204;  
Best Local Similarity 100.0%; Pred. No. 1.e-05;  
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 DEAQSKRGILTLYKYP 15  
|||||  
Db 76 DEAQSKRGILTLYKYP 90

## RESULT 9

US-10-424-599-282905  
; Sequence 282905, Application US/10424599  
; Publication No. US20040031072A1  
; GENERAL INFORMATION:  
; APPLICANT: La Rosa Thomas J  
; APPLICANT: Kovalic David K  
; APPLICANT: Zhou Yihua  
; APPLICANT: Cao Yongwei  
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With  
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement  
; FILE REFERENCE: 38-21(53223)B  
; CURRENT APPLICATION NUMBER: US/10/424,599  
; CURRENT FILING DATE: 2003-04-28  
; NUMBER OF SEQ ID NOS: 285684  
; SEQ ID NO 282905  
; LENGTH: 208  
; TYPE: PRT  
; ORGANISM: Glycine max  
; FEATURE:  
; NAME/KEY: unsure  
; LOCATION: (1)-(208)  
; OTHER INFORMATION: unsure at all Xaa locations  
; FEATURE:  
; OTHER INFORMATION: Clone ID: PAT\_MRT3847\_97486C.1.pep  
US-10-424-599-282905

Query Match 100.0%; Score 76; DB 15; Length 208;  
Best Local Similarity 100.0%; Pred. No. 1.e-05;  
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 DEAQSKRGILTLYKYP 15  
|||||  
Db 58 DEAQSKRGILTLYKYP 72

## RESULT 10

US-10-002-631C-56  
; Sequence 56, Application US/10002631C  
; Publication No. US20030157486A1  
; GENERAL INFORMATION:  
; APPLICANT: Graff, Jonathon M.  
; APPLICANT: Muenster, Matthew  
; TITLE OF INVENTION: METHODS TO IDENTIFY SIGNAL SEQUENCES

; FILE REFERENCE: A34943 090495.0243  
; CURRENT APPLICATION NUMBER: US/10/002,631C  
; CURRENT FILING DATE: 2001-10-31  
; PRIOR APPLICATION NUMBER: 60/300,309  
; PRIOR FILING DATE: 2001-06-21  
; NUMBER OF SEQ ID NOS: 324  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 56  
; LENGTH: 219  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-10-002-631C-56

Query Match 100.0%; Score 76; DB 14; Length 219;  
Best Local Similarity 100.0%; Pred. No. 1.e-05;  
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 DEAQSKRGILTLYKYP 15  
|||||  
Db 83 DEAQSKRGILTLYKYP 97

## RESULT 11

US-10-029-386-32056  
; Sequence 32056, Application US/10029386  
; Publication No. US20030194704A1  
; GENERAL INFORMATION:  
; APPLICANT: Penn, Sharron G.  
; APPLICANT: Hanzel, David K.  
; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR  
; TITLE OF INVENTION: EXPRESSION ANALYSIS TWO  
; FILE REFERENCE: ASOMICA-X-2  
; CURRENT APPLICATION NUMBER: US/10/029,386  
; CURRENT FILING DATE: 2001-12-20  
; NUMBER OF SEQ ID NOS: 34288  
; SOFTWARE: Annotmax Sequence Listing Engine vers. 1.1  
; SEQ ID NO 32056  
; LENGTH: 219  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
; FEATURE:  
; OTHER INFORMATION: MAP TO AC026717.3  
; OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 16  
; OTHER INFORMATION: EXPRESSED IN HELA, SIGNAL = 23  
; OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 6.8  
; OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 20  
; OTHER INFORMATION: EXPRESSED IN HEART, SIGNAL = 9.8  
; OTHER INFORMATION: SWISSPROT HIT: P29751, EVALUATION 3.00e-94  
US-10-029-386-32056

Query Match 100.0%; Score 76; DB 14; Length 219;  
Best Local Similarity 100.0%; Pred. No. 1.e-05;  
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 DEAQSKRGILTLYKYP 15  
|||||  
Db 87 DEAQSKRGILTLYKYP 101

## RESULT 12

US-10-767-701-42770  
; Sequence 42770, Application US/10767701  
; Publication No. US20040172684A1  
; GENERAL INFORMATION:  
; APPLICANT: Kovalic, David K.  
; APPLICANT: Zhou, Yihua  
; APPLICANT: Cao, Yongwei  
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With  
; TITLE OF INVENTION: Plants and Uses Thereof For Plant Improvement  
; FILE REFERENCE: 38-21(53535)B  
; CURRENT APPLICATION NUMBER: US/10/767,701  
; CURRENT FILING DATE: 2004-01-29

```
; NUMBER OF SEQ ID NOS: 63128
; SEQ ID NO 42770
; LENGTH: 230
; TYPE: PRT
; ORGANISM: Sorghum bicolor
; FEATURE:
; OTHER INFORMATION: Clone ID: SORBI-28MAY03-C17410_1.pep
US-10-767-701-42770

Query Match      100.0%; Score 76; DB 16; Length 230;
Best Local Similarity 100.0%; Pred. No. 1.2e-05;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      1 DEAQSKRGILTLYP 15
      |||||
Db      58 DEAQSKRGILTLYP 72

RESULT 13
US-10-424-599-162684
; Sequence 162684, Application US/10424599
; Publication No. US20040031072A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa Thomas J
; APPLICANT: Kovalic David K
; APPLICANT: Zhou Yihua
; APPLICANT: Cao Yongwei
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53223)B
; CURRENT APPLICATION NUMBER: US/10/424,599
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 285684
; SEQ ID NO 162684
; LENGTH: 248
; TYPE: PRT
; ORGANISM: Glycine max
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT3847_117922C.1.pep
US-10-424-599-162684

Query Match      100.0%; Score 76; DB 15; Length 248;
Best Local Similarity 100.0%; Pred. No. 1.3e-05;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      1 DEAQSKRGILTLYP 15
      |||||
Db      58 DEAQSKRGILTLYP 72

RESULT 14
US-10-425-114-59172
; Sequence 59172, Application US/10425114
; Publication No. US2004003488A1
; GENERAL INFORMATION:
; APPLICANT: Liu, Jingdong
; APPLICANT: Zhou, Yihua
; APPLICANT: Kovalic, David K.
; APPLICANT: Screen, Steven E
; APPLICANT: Tabaska, Jack E
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53313)B
; CURRENT APPLICATION NUMBER: US/10/425,114
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 73128
; SEQ ID NO 59172
; LENGTH: 331
; TYPE: PRT
; ORGANISM: Zea mays
; FEATURE:
; OTHER INFORMATION: Clone ID: 700332871_FLI.pep
```

```
US-10-425-114-59172

Query Match      100.0%; Score 76; DB 15; Length 331;
Best Local Similarity 100.0%; Pred. No. 1.8e-05;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      1 DEAQSKRGILTLYP 15
      |||||
Db      12 DEAQSKRGILTLYP 26

RESULT 15
US-10-108-260A-4008
; Sequence 4008, Application US/10108260A
; Publication No. US20040005560A1
; GENERAL INFORMATION:
; APPLICANT: HELIX RESEARCH INSTITUTE
; TITLE OF INVENTION: No. US20040005560A1el full length cDNA
; FILE REFERENCE: H1-A0106
; CURRENT APPLICATION NUMBER: US/10/108,260A
; CURRENT FILING DATE: 2002-03-27
; NUMBER OF SEQ ID NOS: 5458
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 4008
; LENGTH: 342
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-108-260A-4008

Query Match      100.0%; Score 76; DB 15; Length 342;
Best Local Similarity 100.0%; Pred. No. 1.8e-05;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      1 DEAQSKRGILTLYP 15
      |||||
Db      58 DEAQSKRGILTLYP 72

Search completed: April 8, 2005, 12:51:00
Job time : 42.6667 secs
```

**THIS PAGE BLANK (USPTO)**



GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: April 8, 2005, 09:58:56 ; Search time 9.33333 Seconds  
(without alignments)  
154.634 Million cell updates/sec

Title: US-09-423-351C-8  
Perfect score: 75  
Sequence: 1 LASLSTFQQMWSKQ 15  
Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 283416 seqs, 96216763 residues  
Total number of hits satisfying chosen parameters: 283416

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000  
Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : PIR\_79:.\*  
1: pir1.\*  
2: pir2.\*  
3: pir3.\*  
4: pir4.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	75	100.0	42	JE0005	actin alpha, smoot
2	75	100.0	100	S14851	actin - fruit fly
3	75	100.0	137	A28258	actin 5C - fruit f
4	75	100.0	195	S39777	actin beta - pig (
5	75	100.0	213	A61043	actin CAL5 - sea s
6	75	100.0	308	A03000	actin 3 - fruit fl
7	75	100.0	327	S11452	actin (clone 302)
8	75	100.0	349	B25819	actin, fetal skele
9	75	100.0	370	A29664	actin - sea urchin
10	75	100.0	374	ATBOB	actin beta - bovin
11	75	100.0	374	ATBOG	actin gamma - bovi
12	75	100.0	374	JC5818	gamma-actin - huma
13	75	100.0	375	ATBOSM	actin, aortic smoo
14	75	100.0	375	ATRTC	actin beta - rat
15	75	100.0	375	A48324	actin beta, cytoek
16	75	100.0	375	ATCHB	actin beta - chick
17	75	100.0	375	ATHUB	actin beta - human
18	75	100.0	375	ATHUG	actin gamma 1 - hu
19	75	100.0	375	ATMSB	actin beta - mouse
20	75	100.0	375	ATMSG	actin gamma - mous
21	75	100.0	375	ATRBB	actin beta, non-mu
22	75	100.0	375	S11222	actin gamma, cytos
23	75	100.0	375	S33386	actin, cytosolic (
24	75	100.0	375	JT0385	actin gamma - Emer
25	75	100.0	375	S47897	actin 1 - Pneumocy
26	75	100.0	375	T25272	hypothetical prote
27	75	100.0	375	A26836	actin - fission ye
28	75	100.0	375	S71125	actin beta-2, cyto
29	75	100.0	375	S71124	actin beta-1, cyto

30	75	100.0	375	2	S71126	actin beta, cytooso
31	75	100.0	375	2	A55001	actin beta - goose
32	75	100.0	375	2	A54728	actin alpha, card
33	75	100.0	375	2	S42103	actin - Puccinia g
34	75	100.0	376	1	ATPF7	actin 7 - fruit fl
35	75	100.0	376	1	A43552	actin gamma, cytos
36	75	100.0	376	1	ATPF8	actin 8 - fruit fl
37	75	100.0	376	1	ATURS	actin Cy1 - sea ur
38	75	100.0	376	2	A48449	Actin-1A - nematod
39	75	100.0	376	2	S04538	actin 87E - fruit
40	75	100.0	376	2	JC1246	actin - fruit fly
41	75	100.0	376	2	J80189	actin, cytosolic -
42	75	100.0	376	2	J80190	actin, muscle - st
43	75	100.0	376	2	S07288	actin 15A - sea ur
44	75	100.0	376	2	S09578	actin - sea urchin
45	75	100.0	376	2	JQ0154	actin - Hydra atte

ALIGNMENTS

RESULT 1

JE0005  
actin alpha, smooth muscle - rat (fragment)  
C:Species: Rattus norvegicus (Norway rat)  
C:Date: 31-Dec-1991 #sequence\_revision 31-Dec-1991 #text\_change 09-Jul-2004  
C:Accession: JE0005  
R:Hsu, C.Y.J.; Frankel, F.R.  
Gene 69, 345-348, 1988

A:Title: Conserved and unique sequences in the 3'-untranslated region of rat smooth-muscle  
A:Reference number: JE0005; MUID:89172079; PMID:3234770  
A:Accession: JE0005  
A:Molecule type: mRNA  
A:Residues: 1-42 <HSU>  
A:Cross-references: UNIPROT:P70476; GB:M22757; NID:949983; PIDN:AAA74457.1; PID:G959539  
C:Genetics:  
A:Gene: sma  
C:Superfamily: actin  
C:Keywords: muscle; smooth muscle

Query Match 100.0%; Score 75; DB 2; Length 42;  
Best Local Similarity 100.0%; Pred. No. 2.3e-06;  
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 LASLSTFQQMWSKQ 15  
Db 13 LASLSTFQQMWSKQ 27

RESULT 2

S14851  
actin - fruit fly (Drosophila melanogaster)  
C:Species: Drosophila melanogaster  
C:Date: 20-Feb-1995 #sequence\_revision 20-Feb-1995 #text\_change 09-Jul-2004  
C:Accession: S14851  
R:Burn, T.C.; Tobin, S.L.  
submitted to the EMBL Data Library, October 1990  
A:Reference number: S14851  
A:Accession: S14851  
A:Molecule type: DNA  
A:Residues: 1-100 <BUR>  
A:Cross-references: UNIPROT:P02572; EMBL:X54848; NID:g7549; PIDN:CAA38618.1; PID:g7550  
C:Genetics:  
A:Gene: FlyBase:Act42A  
A:Cross-references: FlyBase:FBgn0000043  
C:Superfamily: actin  
C:Keywords: cell motility; cytoskeleton; microfilament; mitosis; structural protein

Query Match 100.0%; Score 75; DB 2; Length 100;  
Best Local Similarity 100.0%; Pred. No. 5.6e-06;  
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 LASLSTFQQMWSKQ 15

```

Db          71 LASLSTFQQMWISKQ 85
|||||
RESULT 3
A28258
actin 5C - fruit fly (Drosophila melanogaster)
C:Species: Drosophila melanogaster
C>Date: 30-Sep-1989 #sequence_revision 30-Sep-1989 #text_change 09-Jul-2004
C:Accession: A28258
R:Vigoreaux, J.O.; Tobin, S.L.
Genes Dev. 1, 1161-1171, 1987
A:Title: Stage-specific selection of alternative transcriptional initiation sites from
A:Reference number: A28258; MUID:88112795; PMID:312314
A:Accession: A28258
A:Molecule type: mRNA
A:Residues: 1-137 <VIG>
A:CROSS-references: UNIPROT:P10987
A:Note: the authors translated the codon GAG for residue 96 as Gly
C:Genetics:
A:Gene: FlyBase:Act5C
A:CROSS-references: FlyBase:FBgn0000042
C:Superfamily: actin
C:Keywords: methylated amino acid
F:74/Modified site: 3'-methylhistidine (His) #status predicted

Query Match          100.0%; Score 75; DB 2; Length 137;
Best Local Similarity 100.0%; Pred. No. 7.6e-06;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY          1 LASLSTFQQMWISKQ 15
|||||
Db          108 LASLSTFQQMWISKQ 122
|||||
RESULT 4
S39777
actin beta - pig (fragments)
C:Species: Sus scrofa domestica (domestic pig)
C>Date: 07-Oct-1994 #sequence_revision 12-Apr-1996 #text_change 09-Jul-2004
C:Accession: S39777
R:Adamski, J.; Husen, B.; Thole, H.H.; Groeschel-Stewart, U.; Jungblut, P.W.
Biochem. J. 296, 797-802, 1993
A:Title: Linkage of 17-beta-estradiol dehydrogenase to actin by epsilon-(gamma-glutamyl)
A:Reference number: S39777; MUID:94107247; PMID:8280079
A:Accession: S39777
A:Molecule type: protein
A:Residues: 1-20;21-52;53-79;80-85;86-95;96-116;117-151;152-159;160-195 <ADA>
A:CROSS-references: UNIPROT:Q7M3B0
C:Superfamily: actin

Query Match          100.0%; Score 75; DB 2; Length 195;
Best Local Similarity 100.0%; Pred. No. 1.1e-05;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY          1 LASLSTFQQMWISKQ 15
|||||
Db          168 LASLSTFQQMWISKQ 182
|||||
RESULT 5
A61043
actin CM15 - sea squirt (Styela clava) (fragments)
C:Species: Styela clava
C>Date: 31-Dec-1993 #sequence_revision 31-Dec-1993 #text_change 09-Jul-2004
C:Accession: A61043
R:Beach, R.L.; Jeffery, W.R.
Dev. Genet. 11, 2-14, 1990
A:Title: Temporal and spatial expression of a cytoskeletal actin gene in the ascidian
A:Reference number: A61043; MUID:90298580; PMID:2361333
A:Accession: A61043
A>Status: not compared with conceptual translation
A:Molecule type: mRNA

```

```

A:Residues: 1-213 <BEA>
A:CROSS-references: UNIPROT:Q7M3Y7
C:Comment: This sequence is expressed in cells undergoing rapid cell division.
C:Superfamily: actin
C:Keywords: cytoskeleton; methylated amino acid; mitosis; structural protein
F:73/Modified site: 3'-methylhistidine (His) #status predicted

Query Match          100.0%; Score 75; DB 2; Length 213;
Best Local Similarity 100.0%; Pred. No. 1.2e-05;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY          1 LASLSTFQQMWISKQ 15
|||||
Db          184 LASLSTFQQMWISKQ 198
|||||
RESULT 6
A03000
actin 3 - fruit fly (Drosophila melanogaster) (fragments)
C:Species: Drosophila melanogaster
C>Date: 31-Mar-1991 #sequence_revision 31-Mar-1991 #text_change 09-Jul-2004
C:Accession: A03000
R:Fyrberg, B.A.; Bond, B.J.; Hershey, N.D.; Mixter, K.S.; Davidson, N.
Cell 24, 107-116, 1981
A:Title: The actin genes of Drosophila: protein coding regions are highly conserved but
A:Reference number: A03000; MUID:81210174; PMID:6263481
A:Accession: A03000
A:Molecule type: DNA
A:Residues: 1-308 <FYR>
A:CROSS-references: UNIPROT:P02572
A:Note: there are 68 unsequenced residues between positions 160 and 161. Partial sequence
C:Genetics:
A:Gene: FlyBase:Act42A
A:CROSS-references: FlyBase:FBgn0000043
A:Map position: 42A
C:Superfamily: actin
C:Keywords: methylated amino acid
F:74/Modified site: 3'-methylhistidine (His) #status predicted

Query Match          100.0%; Score 75; DB 2; Length 308;
Best Local Similarity 100.0%; Pred. No. 1.7e-05;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY          1 LASLSTFQQMWISKQ 15
|||||
Db          279 LASLSTFQQMWISKQ 293
|||||
RESULT 7
S11452
actin (clone 302) - brine shrimp (fragment)
C:Species: Artemia sp. (brine shrimp)
C>Date: 17-Apr-1993 #sequence_revision 17-Apr-1993 #text_change 09-Jul-2004
C:Accession: S11452
R:Macias, M.T.; Sastre, L.
Nucleic Acids Res. 19, 5219-5225, 1990
A:Title: Molecular cloning and expression of four actin isoforms during Artemia developm
A:Reference number: S11450; MUID:90384823; PMID:2402445
A:Accession: S11452
A:Molecule type: mRNA
A:Residues: 1-327 <MAC>
A:CROSS-references: UNIPROT:P18602; EMBL:X52604; NID:g5666; PIDN:CAA36837.1; PID:g829162
C:Superfamily: actin
C:Keywords: methylated amino acid; structural protein
F:75/Modified site: 3'-methylhistidine (His) #status predicted

Query Match          100.0%; Score 75; DB 2; Length 327;
Best Local Similarity 100.0%; Pred. No. 1.8e-05;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY          1 LASLSTFQQMWISKQ 15
|||||

```

Db 298 LASLSTFOQMWSKQ 312

RESULT 8

B25819  
actin, fetal skeletal/adult cardiac muscle - mouse (fragment)  
C:Species: Mus musculus (house mouse)  
C>Date: 23-Aug-1987 #sequence\_revision 23-Aug-1987 #text\_change 09-Jul-2004  
C:Accession: B25819  
R:Alonso, S.; Minty, A.; Bourlet, Y.; Buckingham, M.  
J. Mol. Evol. 23, 11-22, 1986  
A>Title: Comparison of three actin-coding sequences in the mouse; evolutionary relations  
A:Reference number: A25819; MUID:86200234; PMID:3084797  
A:Accession: B25819  
A:Molecule type: mRNA  
A:Residues: 1-349 <ALO>  
A:Cross-references: UNIPROT:P061275; GB:J00381; GB:X03767; GB:M10652; NID:G49869; PIDN:CA  
C:Superfamily: actin  
C:Keywords: cardiac muscle; heart; methylated amino acid; muscle; skeletal muscle  
F:47/Modified site: 3'-methylhistidine (His) #status predicted

Query Match 100.0%; Score 75; DB 2; Length 349;  
Best Local Similarity 100.0%; Pred. No. 2e-05;  
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 LASLSTFOQMWSKQ 15

Db 320 LASLSTFOQMWSKQ 334

RESULT 9

A29664  
actin - sea urchin (Strongylocentrotus purpuratus)  
C:Species: Strongylocentrotus purpuratus (purple urchin)  
C>Date: 15-Dec-1988 #sequence\_revision 30-Sep-1991 #text\_change 05-Dec-1997  
C:Accession: A29664  
R:Crain Jr., W.R.; Boshar, M.F.; Cooper, A.D.; Durica, D.S.; Nagy, A.; Steffen, D.  
J. Mol. Evol. 25, 37-45, 1987  
A>Title: The sequence of a sea urchin muscle actin gene suggests a gene conversion with  
A:Reference number: A29664; MUID:87311761; PMID:3114500  
A:Accession: A29664  
A:Molecule type: DNA  
A:Residues: 1-370 <CRA>  
A:Cross-references: GB:X05739; GB:X05740; GB:X05741; GB:X05742; GB:X05743  
A:Note: the authors translated the codon CAG for residue 260 as Glu  
C:Superfamily: actin  
C:Keywords: methylated amino acid  
F:73/Modified site: 3'-methylhistidine (His) #status predicted

Query Match 100.0%; Score 75; DB 2; Length 370;  
Best Local Similarity 100.0%; Pred. No. 2.1e-05;  
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 LASLSTFOQMWSKQ 15

Db 343 LASLSTFOQMWSKQ 357

RESULT 10

ATBOG  
actin beta - bovine (tentative sequence)  
C:Species: Bos primigenius taurus (cattle)  
C>Date: 30-Sep-1991 #sequence\_revision 30-Sep-1991 #text\_change 09-Jul-2004  
C:Accession: E14185; A39105; A02999; A14185  
R:Vandekerckhove, J.; Weber, K.  
Eur. J. Biochem. 90, 451-462, 1978  
A>Title: Actin amino-acid sequences. Comparison of actins from calf thymus, bovine brain  
A:Reference number: A14185; MUID:79045349; PMID:213279  
A:Accession: E14185  
A:Molecule type: protein  
A:Residues: 1-374 <VAN>  
A:Cross-references: UNIPROT:P60712  
A:Note: only peptides that differed in composition from the corresponding peptides of ra

R;Degen, J.L.; Neubauer, M.G.; Degen, S.J.F.; Seyfried, C.E.; Morris, D.R.  
J. Biol. Chem. 258, 12153-12162, 1983  
A>Title: Regulation of protein synthesis in mitogen-activated bovine lymphocytes. Analyt-  
A:Reference number: A39105; MUID:84032385; PMID:6195151  
A:Accession: A39105  
A:Molecule type: mRNA  
A:Residues: 76-227;344-374 <DEG>  
A:Cross-references: GB:K00622; GB:K00623  
A:Note: actins beta and gamma were not distinguished in this study  
C:Comment: Vertebrate nonmuscle cells contain two highly conserved cytoskeletal actins,  
C:Superfamily: actin  
C:Keywords: blocked amino end; cell motility; cytoskeleton; methylated amino acid; micr-  
F:71/Modified site: blocked amino end (Asp) (probably acetylated) #status experimental  
F:72/Modified site: 3'-methylhistidine (His) #status predicted

Query Match 100.0%; Score 75; DB 1; Length 374;  
Best Local Similarity 100.0%; Pred. No. 2.1e-05;  
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 LASLSTFOQMWSKQ 15

Db 345 LASLSTFOQMWSKQ 359

RESULT 11

ATBOG  
actin gamma - bovine (tentative sequence)  
C:Species: Bos primigenius taurus (cattle)  
C>Date: 30-Sep-1991 #sequence\_revision 30-Sep-1991 #text\_change 09-Jul-2004  
C:Accession: B14185; A02999  
R:Vandekerckhove, J.; Weber, K.  
Eur. J. Biochem. 90, 451-462, 1978  
A>Title: Actin amino-acid sequences. Comparison of actins from calf thymus, bovine brain  
A:Reference number: A14185; MUID:79045349; PMID:213279  
A:Accession: B14185  
A:Molecule type: protein  
A:Residues: 1-374 <VAN>  
A:Cross-references: UNIPROT:P02571  
A:Note: only peptides that differed in composition from the corresponding peptides of r  
C:Comment: Vertebrate nonmuscle cells contain two highly conserved cytoskeletal actins,  
C:Superfamily: actin  
C:Keywords: blocked amino end; cell motility; cytoskeleton; methylated amino acid; micr-  
F:71/Modified site: blocked amino end (Glu) (probably acetylated) #status experimental  
F:72/Modified site: 3'-methylhistidine (His) #status predicted

Query Match 100.0%; Score 75; DB 1; Length 374;  
Best Local Similarity 100.0%; Pred. No. 2.1e-05;  
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 LASLSTFOQMWSKQ 15

Db 345 LASLSTFOQMWSKQ 359

RESULT 12

JC5818  
gamma-actin - human  
C:Species: Homo sapiens (man)  
C>Date: 20-Feb-1998 #sequence\_revision 13-Mar-1998 #text\_change 09-Jul-2004  
C:Accession: JC5818; PC4501  
R:Hauschildt, S.; Schwarz, C.; Heine, H.; Ulmer, A.J.; Flad, H.D.; Rietschel, E.T.; Jen  
Biochem. Biophys. Res. Commun. 241, 670-674, 1997  
A>Title: Actin: A target of lipopolysaccharide-induced phosphorylation in human monocyte  
A:Reference number: JC5818; MUID:98096379; PMID:9434766  
A:Accession: JC5818  
A:Molecule type: protein  
A:Residues: 1-374 <HAU>  
A:Cross-references: UNIPROT:P02571  
A:Experimental source: monocyte  
A:Accession: PC4501  
A:Molecule type: protein  
A:Residues: 1-61;84-112;147-190;196-209;215-253;335-358 <HA2>  
A:Experimental source: monocyte

C;Comment: This protein is involved in a signal transduction that eventually leads to md  
C;Superfamily: actin

Query Match 100.0%; Score 75; DB 2; Length 374;  
Best Local Similarity 100.0%; Pred. No. 2.1e-05;  
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 LASLSTFQQMWISKQ 15  
|||||

Db 345 LASLSTFQQMWISKQ 359

## RESULT 13

ATBOSM

C;Species: Bos primigenius taurus (cattle)  
C;Date: 30-Jun-1987 #sequence\_revision 30-Jun-1987 #text\_change 06-Sep-1996  
C;Accession: A02997; S13480  
R;Vandekerckhove, J.; Weber, K.  
A;Title: The complete amino acid sequence of actins from bovine aorta, bovine heart, bovine skeletal muscle, and bovine brain.

A;Reference number: A02997; MUID:80047657; PMID:499690

A;Accession: A02997

A;Molecule type: protein

A;Residues: 1-375 <VAN>

R;Zevgoulis, V.G.; Sotiropoulos, T.G.; Evangelopoulos, A.E.

Biochim. Biophys. Acta 1091, 222-230, 1991

A;Title: Phosphorylase kinase from bovine stomach smooth muscle: a Ca (2+)-dependent protein with a high affinity for ATP.

A;Reference number: S13480; MUID:91137633; PMID:1995080

A;Accession: S13480

A;Molecule type: protein

A;Residues: 40-49 <ZEV>

A;Experimental source: stomach

A;Note: this material appears to be actin of aortic smooth muscle type or a related molecule.

C;Superfamily: actin

C;Keywords: acetylated amino end; methylated amino acid; muscle contraction

F;1/Modified site: acetylated amino end (Glu) #status predicted

F;73/Modified site: 3'-methylhistidine (His) #status experimental

Query Match 100.0%; Score 75; DB 1; Length 375;

Best Local Similarity 100.0%; Pred. No. 2.1e-05;

Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 LASLSTFQQMWISKQ 15  
|||||

Db 346 LASLSTFQQMWISKQ 360

## RESULT 14

ATRC

C;Species: Rattus norvegicus (Norway rat)  
C;Date: 13-Jun-1983 #sequence\_revision 13-Jun-1983 #text\_change 22-Jun-1999  
C;Accession: A38571; A02999

R;Nudel, U.; Zakut, R.; Shani, M.; Neuman, S.; Levy, Z.; Yaffe, D.

Nucleic Acids Res. 11, 1759-1771, 1983

A;Title: The nucleotide sequence of the rat cytoplasmic beta-actin gene.

A;Reference number: A38571; MUID:83168920; PMID:6300777

A;Accession: A38571

A;Molecule type: DNA

A;Residues: 1-375 <NUD>

A;Cross-references: GB:J00691; NID:g202653; PIDN:AAA40657.1; PID:g202654

C;Comment: Vertebrate nonmuscle cells contain two highly conserved cytoskeletal actins,

C;Genetics:

A;Introns: 41/3; 121/3; 268/1; 328/3

C;Superfamily: actin

C;Keywords: cell motility; cytoskeleton; methylated amino acid; microfilament; mitosis;

F;2-375/Product: actin beta #status predicted <NAI>

F;73/Modified site: 3'-methylhistidine (His) #status predicted

Query Match 100.0%; Score 75; DB 1; Length 375;

Best Local Similarity 100.0%; Pred. No. 2.1e-05;

Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 LASLSTFQQMWISKQ 15  
|||||

Db 346 LASLSTFQQMWISKQ 360

## RESULT 15

A48324

actin beta, cytoskeletal - common carp

C;Species: Cyprinus carpio (common carp)

C;Date: 03-Feb-1994 #sequence\_revision 11-Apr-1997 #text\_change 09-Jul-2004

C;Accession: A48324

R;Liu, Z.; Zhu, Z.; Roberg, K.; Faras, A.; Guise, K.; Kapuscinski, A.R.; Hackett, P.B.

DNA Seq. 1, 125-136, 1990

A;Title: Isolation and characterization of beta-actin gene of carp (Cyprinus carpio).

A;Reference number: A48324; MUID:92190540; PMID:2134183

A;Accession: A48324

A;Molecule type: DNA

A;Residues: 1-375 <LIU>

A;Cross-references: UNIPROT:P83750; GB:M24113; NID:g213041; PIDN:AAA68886.1; PID:g213042

A;Note: the authors translated the codon TTC for residue 21 as Pro, AAG for residue 50 as

7 as Pro

A;Note: the authors failed to translated the codon GGT for residue 42 as Gly

C;Comment: Vertebrate nonmuscle cells contain two highly conserved cytoskeletal actins,

C;Genetics: 41/3; 121/3; 268/3; 328/3

C;Superfamily: actin

C;Keywords: acetylated amino end; cell motility; cytoskeleton; methylated amino acid; mi

F;2-375/Product: actin beta, cytoskeletal #status predicted <MAT>

F;73/Modified site: acetylated amino end (Asp) (in mature form) #status predicted

F;73/Modified site: 3'-methylhistidine (His) #status predicted

Query Match 100.0%; Score 75; DB 1; Length 375;

Best Local Similarity 100.0%; Pred. No. 2.1e-05;

Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 LASLSTFQQMWISKQ 15  
|||||

Db 346 LASLSTFQQMWISKQ 360

Search completed: April 8, 2005, 10:53:06

Job time : 9.33333 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: April 8, 2005, 09:58:56 ; Search time 57.8667 Seconds  
(without alignments)  
132.739 Million cell updates/sec

Title: US-09-423-351C-8  
Perfect score: 75  
Sequence: 1 LASLSTFOQMWSKQ 15

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 1612378 seqs, 512079187 residues

Total number of hits satisfying chosen parameters: 1612378

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Uniprot\_03:\*  
1: uniprot\_sprot:\*  
2: uniprot\_trembl:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	75	100.0	33	1 ACT DICVI	Q24733 dictyocaulu
2	75	100.0	38	2 Q6VPP6	Q6VPP6 ixodes rici
3	75	100.0	45	2 Q6J0R3	Q6J0R3 hydra atten
4	75	100.0	45	2 Q6J0R4	Q6J0R4 hydra atten
5	75	100.0	50	2 Q9DDZ2	Q9DDZ2 micropterus
6	75	100.0	55	2 Q8AY69	Q8AY69 oncorhynch
7	75	100.0	56	2 Q57575	Q57575 cynops pyr
8	75	100.0	93	2 Q7ZZZ0	Q7ZZZ0 anas platyr
9	75	100.0	96	2 Q90X51	Q90X51 fundulus he
10	75	100.0	100	2 Q9TYD9	Q9TYD9 styela clav
11	75	100.0	106	2 Q80UA7	Q80UA7 marmota mon
12	75	100.0	113	2 Q6Y260	Q6Y260 pagrus majo
13	75	100.0	123	2 Q6UIS1	Q6UIS1 macaca mula
14	75	100.0	123	2 Q8UIS2	Q8UIS2 pan troglod
15	75	100.0	124	2 Q9DFK3	Q9DFK3 gillichthys
16	75	100.0	131	2 Q9XSB1	Q9XSB1 tupaja glis
17	75	100.0	134	2 Q99NC5	Q99NC5 mus muscul
18	75	100.0	134	2 Q99NC6	Q99NC6 rattus norv
19	75	100.0	149	2 Q6PTC1	Q6PTC1 sitobion av
20	75	100.0	149	2 Q6PTC2	Q6PTC2 sitobion av
21	75	100.0	154	1 ACT4_LYTP1	Q25380 lytechinus
22	75	100.0	159	2 Q96DE1	Q96DE1 homo sapien
23	75	100.0	165	2 Q96FU6	Q96FU6 homo sapien
24	75	100.0	172	1 ACT3_LYTP1	Q25379 lytechinus
25	75	100.0	172	1 ACTM_LYTP1	Q25381 lytechinus
26	75	100.0	181	2 Q6XJ33	Q6XJ33 drosophila
27	75	100.0	195	2 Q7M3B0	Q7M3B0 sus scrofa
28	75	100.0	213	2 Q7M3Y7	Q7M3Y7 styela clav
29	75	100.0	243	2 Q8MVP7	Q8MVP7 boltenia vi
30	75	100.0	244	2 Q94FN6	Q94FN6 lotharella
31	75	100.0	257	2 Q9P443	Q9P443 aspergillus

Q6jxi0 hedriocysti  
Q6jxi1 clathrulina  
Q6jxi2 raphidlophr  
Q6jxi3 homo sapien  
Q6jxi4 ophiostoma  
Q6jxi5 ophiostoma  
Q6jxi6 ophiostoma  
Q6jxi7 cricetidae  
Q6jxi8 lama glama  
Q6jxi9 heliocidari  
Q6jxi10 heliocidari  
Q6jxi11 culicoides  
Q6jxi12 mustela put  
Q6jxi13 littorina l  
Q6jxi14 artemia sp.  
Q6jxi15 debaryomyce

#### ALIGNMENTS

##### RESULT 1

ACT\_DICVI STANDARD; PRT; 33 AA.  
AC Q24733;  
DT 01-NOV-1997 (Rel. 35, Created)  
DT 01-NOV-1997 (Rel. 35, Last sequence update)  
DT 15-JUL-1999 (Rel. 38, Last annotation update)  
DE Actin (Fragment).  
OS Dictyocaulus viviparus (Bovine lungworm).  
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Strongylida;  
OC Trichostrongyloidea; Dictyocaulidae; Dictyocaulinae; Dictyocaulus.  
OX NCBI\_TaxID=29172;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC TISSUE=Larva;  
RA von Samson-Hammelstjerna G., Wunderlich G., Muehlethlegel F.,  
RA Froesch M., Schnieder T.;  
RL Submitted (DEC-1995) to the EMBL/GenBank/DBJ databases.  
CC -!- FUNCTION: Actins are highly conserved proteins that are involved  
CC in various types of cell motility and are ubiquitously expressed  
CC in all eukaryotic cells.  
CC -!- SUBCELLULAR LOCATION: Cytoplasmic.  
CC -!- SIMILARITY: Belongs to the actin family.  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/  
CC or send an email to license@isb-sib.ch).  
CC -----  
CC EMBL; U40590; AAA83551.1; -  
CC InterPro; IPR004001; Actin.  
CC InterPro; IPR004000; Actin\_like.  
CC Pfam; PF00022; Actin; 1.  
CC PROSITE; PS00406; ACTINS\_1; PARTIAL.  
CC PROSITE; PS00432; ACTINS\_2; 1.  
CC PROSITE; PS01132; ACTINS\_ACT\_LIKE; PARTIAL.  
CC Structural protein.  
CC NON TER 1  
SQ SEQUENCE 33 AA; 3830 MW; 72E2DB4D5D5CDAEC CRC64;

Query Match 100.0%; Score 75; DB 1; Length 33;  
Best Local Similarity 100.0%; Pred. No. 8.4e-06;  
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 LASLSTFOQMWSKQ 15

DB 4 LASLSTFOQMWSKQ 18

##### RESULT 2

Q6VPP6

```

ID Q6VPP6 PRELIMINARY; PRT; 38 AA.
AC Q6VPP6;
DT 05-JUL-2004 (TrEMBLrel. 27, Created)
DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
DE Gamma actin-like protein (Fragment).
OS Ixodes ricinus (Sheep tick).
OC Eukaryota; Metazoa; Arthropoda; Chelicerata; Arachnida; Acari;
OC Parasitiformes; Ixodida; Ixodidae; Ixodes.
OX NCBI_TaxID=34613;
RN [1]
RP SEQUENCE FROM N.A.
RA Rudenko N.K., Golovchenko M.P., Edwards M.J., Grubhoffer L.;
RL Submitted (JUL-2003) to the EMBL/GenBank/DBJ databases.
DR EMBL; AY333957; AAQ23081.1; -.
DR GO; GO:0005884; C:actin filament; IEA.
DR GO; GO:0003774; F:motor activity; IEA.
DR GO; GO:0005200; F:structural constituent of cytoskeleton; IEA.
DR InterPro; IPR004001; Actin.
DR PROSITE; PS00432; ACTINS_2; 1.
FT NON_TER 1
SQ SEQUENCE 38 AA; 4331 MW; D52BA0C57DC38864 CRC64;

Query Match 100.0%; Score 75; DB 2; Length 38;
Best Local Similarity 100.0%; Pred. No. 9.7e-06;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 LASLSTFQQMWSKQ 15
Db 9 LASLSTFQQMWSKQ 23

RESULT 3
ID Q6J0R3 PRELIMINARY; PRT; 45 AA.
AC Q6J0R3;
DT 05-JUL-2004 (TrEMBLrel. 27, Created)
DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
DE Actin 1.7 (Fragment).
OS Hydra attenuata (Hydra) (Hydra vulgaris).
OC Eukaryota; Metazoa; Cnidaria; Hydrozoa; Hydroidea; Anthomedusae;
OC Hydridae; Hydra.
OX NCBI_TaxID=6087;
RN [1]
RP SEQUENCE FROM N.A.
RA Fisher D.A., Bode H.R.;
RL Submitted (APR-2004) to the EMBL/GenBank/DBJ databases.
DR EMBL; AY609159; AAT37544.1; -.
DR GO; GO:0005884; C:actin filament; IEA.
DR GO; GO:0003774; F:motor activity; IEA.
DR GO; GO:0005200; F:structural constituent of cytoskeleton; IEA.
DR InterPro; IPR004001; Actin.
DR PROSITE; PS00432; ACTINS_2; 1.
FT NON_TER 1
SQ SEQUENCE 45 AA; 5159 MW; 7796DEF78D574D0 CRC64;

Query Match 100.0%; Score 75; DB 2; Length 45;
Best Local Similarity 100.0%; Pred. No. 1.2e-05;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 LASLSTFQQMWSKQ 15
Db 16 LASLSTFQQMWSKQ 30

RESULT 4
ID Q6J0R4 PRELIMINARY; PRT; 45 AA.
AC Q6J0R4;
DT 05-JUL-2004 (TrEMBLrel. 27, Created)
DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
DE Actin 1.6 (Fragment).
OS Hydra attenuata (Hydra) (Hydra vulgaris).
OC Eukaryota; Metazoa; Cnidaria; Hydrozoa; Hydroidea; Anthomedusae;
OC Hydridae; Hydra.
OX NCBI_TaxID=6087;
RN [1]
RP SEQUENCE FROM N.A.
RA Fisher D.A., Bode H.R.;
RL Submitted (APR-2004) to the EMBL/GenBank/DBJ databases.
DR EMBL; AY609159; AAT37544.1; -.
DR GO; GO:0005884; C:actin filament; IEA.
DR GO; GO:0003774; F:motor activity; IEA.
DR GO; GO:0005200; F:structural constituent of cytoskeleton; IEA.
DR InterPro; IPR004001; Actin.
DR PROSITE; PS00432; ACTINS_2; 1.
FT NON_TER 1
SQ SEQUENCE 45 AA; 5159 MW; 7796DEF78D574D0 CRC64;

Query Match 100.0%; Score 75; DB 2; Length 45;
Best Local Similarity 100.0%; Pred. No. 1.2e-05;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 LASLSTFQQMWSKQ 15
Db 16 LASLSTFQQMWSKQ 30

RESULT 5
ID Q9DDZ2 PRELIMINARY; PRT; 50 AA.
AC Q9DDZ2;
DT 01-MAR-2001 (TrEMBLrel. 16, Created)
DT 01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
DE Beta-actin (Fragment).
OS Micropterus salmoides (Largemouth bass).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;
OC Acanthomorpha; Acanthopterygii; Percomorpha; Perciformes; Percoidae;
OC Centrarchidae; Micropterus.
OX NCBI_TaxID=27706;
RN [1]
RP SEQUENCE FROM N.A.
RA Bowman C.J., Sabo-Attwood T.L., Lee H.S., Denalow N.D.;
RL Submitted (APR-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF253319; AAG44624.1; -.
DR HSSP; P02568; IMA9.
DR GO; GO:0015629; C:actin cytoskeleton; IEA.
DR GO; GO:0005200; F:structural constituent of cytoskeleton; IEA.
DR InterPro; IPR004000; Actin_like.
DR Pfam; PF00022; Actin; 1.
FT NON_TER 1
SQ SEQUENCE 50 AA; 5694 MW; 44542890573899BB CRC64;
```

Query Match 100.0%; Score 75; DB 2; Length 50;  
Best Local Similarity 100.0%; Pred. No. 1.3e-05;  
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 LASLSTFOQMWSKQ 15  
DB 35 LASLSTFOQMWSKQ 49

RESULT 6

Q8AY69 ID Q8AY69 PRELIMINARY; PRT; 55 AA.  
AC Q8AY69;  
DT 01-MAR-2003 (TREMBlrel. 23, Created)  
DT 01-MAR-2003 (TREMBlrel. 23, Last sequence update)  
DT 01-OCT-2003 (TREMBlrel. 25, Last annotation update)  
DE Beta-actin (Fragment).  
OS Oncorhynchus mykiss (Rainbow trout) (Salmo gairdneri).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Actinopterygii; Neopterygii; Teleostei; Euteleostei;  
OC Protacanthopterygii; Salmoniformes; Salmonidae; Oncorhynchus.  
OC NCBI\_TaxID=8022;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC TISSUE=Brain;  
RA Schindt A.R., Alonso M., Leong J.C., Schreck C.B.;  
RL Submitted (OCT-2002) to the EMBL/GenBank/DBJ databases.  
RL EMBL; AF550583; AAN74801.1; --  
DR HSSP; P02568; 1MA9.  
DR GO; GO:0015629; C:actin cytoskeleton; IEA.  
DR GO; GO:0005200; F:structural constituent of cytoskeleton; IEA.  
DR InterPro; IPR004000; Actin\_like.  
DR Pfam; PF00022; Actin; 1.  
FT NON\_TER 1  
FT NON\_TER 55  
SQ SEQUENCE 55 AA; 6163 MW; 944319FB34DD79D8 CRC64;

Query Match 100.0%; Score 75; DB 2; Length 55;  
Best Local Similarity 100.0%; Pred. No. 1.4e-05;  
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 LASLSTFOQMWSKQ 15  
DB 40 LASLSTFOQMWSKQ 54

RESULT 7

O57575 ID O57575 PRELIMINARY; PRT; 56 AA.  
AC O57575;  
DT 01-JUN-1998 (TREMBlrel. 06, Created)  
DT 01-JUN-1998 (TREMBlrel. 06, Last sequence update)  
DT 01-JUN-2003 (TREMBlrel. 24, Last annotation update)  
DE Alpha cardiac actin (fragment).  
OS Cynops pyrrhogaster (Japanese common newt).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Amphibia; Batrachia; Caudata; Salamandroidea; Salamandridae; Cynops.  
OC NCBI\_TaxID=8330;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC Mizuno M., Takabatake T., Takahashi T.C., Takashima K.;  
RA "pax-6 gene expression in newt eye development.";  
RL Dev. Genes Evol. 207:167-176(1997).  
DR EMBL; D64037; BAA24031.1; --  
DR HSSP; P02568; LLCU.  
DR GO; GO:0005884; C:actin filament; IEA.  
DR GO; GO:0003774; F:motor activity; IEA.  
DR GO; GO:0005200; F:structural constituent of cytoskeleton; IEA.  
DR InterPro; IPR004001; Actin.  
DR InterPro; IPR004000; Actin\_like.  
DR Pfam; PF00022; Actin; 1.  
DR PROSITE; PS00432; ACTINS\_1.

Query Match 100.0%; Score 75; DB 2; Length 93;  
Best Local Similarity 100.0%; Pred. No. 2.4e-05;  
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 LASLSTFOQMWSKQ 15  
DB 64 LASLSTFOQMWSKQ 78

RESULT 9

Q90X51 ID Q90X51 PRELIMINARY; PRT; 96 AA.  
AC Q90X51;  
DT 01-DEC-2001 (TREMBlrel. 19, Created)  
DT 01-DEC-2001 (TREMBlrel. 19, Last sequence update)  
DT 01-JUN-2003 (TREMBlrel. 24, Last annotation update)  
DE Beta-actin (Fragment).  
OS Fundulus heteroclitus (Killifish) (Mummichog).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;  
OC Acanthomorpha; Acanthopterygii; Percormorpha; Atherinomorpha;  
OC Cyprinodontiformes; Fundulidae; Fundulus.  
OC NCBI\_TaxID=8078;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC TISSUE=Liver;  
RA Rolling J.A., Baldwin W.S.;  
RL Submitted (OCT-2001) to the EMBL/GenBank/DBJ databases.  
RL EMBL; AF435092; AAL29465.1; --  
DR HSSP; P02568; 1QZ5.  
DR GO; GO:0005884; C:actin filament; IEA.  
DR GO; GO:0003774; F:motor activity; IEA.

DR GO:0005200; F:structural constituent of cytoskeleton; IEA.  
DR Pfam: PF00022; Actin; 1.  
DR PROSITE: PS00432; ACTINS\_2; 1.  
FT NON\_TER 1  
SQ SEQUENCE 96 AA; 10798 MW; 7440DA6AE5A070C4 CRC64;

Query Match 100.0%; Score 75; DB 2; Length 96;  
Best Local Similarity 100.0%; Pred. No. 2.5e-05; Indels 0; Gaps 0;  
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
Qy 1 LASLSTFQQMWSKQ 15  
Db 78 LASLSTFQQMWSKQ 92

RESULT 10  
Q9TYD9 PRELIMINARY; PRT; 100 AA.  
ID Q9TYD9; AC Q9TYD9; DT 01-MAY-2000 (TREMBlrel. 13, Created)  
DT 01-MAY-2000 (TREMBlrel. 13, Last sequence update)  
DT 01-JUN-2003 (TREMBlrel. 24, Last annotation update)  
DE Alpha-muscle actin (Fragment).  
GN Name=TB24;  
OS Styela clava (sea squirt).  
OC Eukaryota; Metazoa; Chordata; Urochordata; Ascidiacea;  
OC Stolidobranchia; Styelidae; Styela.  
OX NCBI\_TaxID=7725;  
RN [1]  
RP MEDLINE=92249656; PubMed=1577198;  
RA Beach R.L., Jeffery W.R.;  
RT "Multiple actin genes encoding the same alpha-muscle isoform are expressed during ascidian development.";  
RL Dev. Biol. 151:55-66 (1992).  
DR EMBL; L13788; AAA29846.1; -.  
DR HSSP; P02568; 1Q25.  
DR GO:0005884; C:actin filament; IEA.  
DR GO:0003774; F:motor activity; IEA.  
DR GO:0005200; F:structural constituent of cytoskeleton; IEA.  
DR InterPro; IPR004001; Actin.  
DR InterPro; IPR004000; Actin\_like.  
DR Pfam; PF00022; Actin; 1.  
DR PROSITE; PS00432; ACTINS\_2; 1.  
FT NON\_TER 1  
SQ SEQUENCE 100 AA; 11264 MW; 108CE34548BABB5E CRC64;

Query Match 100.0%; Score 75; DB 2; Length 100;  
Best Local Similarity 100.0%; Pred. No. 2.6e-05; Indels 0; Gaps 0;  
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
Qy 1 LASLSTFQQMWSKQ 15  
Db 71 LASLSTFQQMWSKQ 85

RESULT 11  
Q80UA7 PRELIMINARY; PRT; 106 AA.  
ID Q80UA7; AC Q80UA7; DT 01-JUN-2003 (TREMBlrel. 24, Created)  
DT 01-JUN-2003 (TREMBlrel. 24, Last sequence update)  
DT 01-MAR-2004 (TREMBlrel. 26, Last annotation update)  
DE Beta-actin (Fragment).  
OS Marmota monax (woodchuck).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Sciuridae; Sciurinae;  
OC Marmota.  
OX NCBI\_TaxID=9995;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=22554704; PubMed=12668969; DOI=10.1053/jhep.2003.50154;

RA Wang Y., Menne S., Jacob J.R., Tennant B.C., Gerin J.L., Cote P.J.;  
RT "Role of type 1 versus type 2 immune responses in liver during the onset of chronic woodchuck hepatitis virus infection.";  
RL Hepatology 37:771-780 (2003).  
DR EMBL; AY170121; AAC39434.1; -.  
DR HSSP; P10983; 1D4X.  
DR GO:0005884; C:actin filament; IEA.  
DR GO:0003774; F:motor activity; IEA.  
DR GO:0005200; F:structural constituent of cytoskeleton; IEA.  
DR InterPro; IPR004001; Actin.  
DR InterPro; IPR004000; Actin\_like.  
DR Pfam; PF00022; Actin; 1.  
DR PROSITE; PS00432; ACTINS\_2; 1.  
FT NON\_TER 1  
SQ SEQUENCE 106 AA; 11864 MW; 9FFDC145651FE343 CRC64;

Query Match 100.0%; Score 75; DB 2; Length 106;  
Best Local Similarity 100.0%; Pred. No. 2.8e-05; Indels 0; Gaps 0;  
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 LASLSTFQQMWSKQ 15  
Db 77 LASLSTFQQMWSKQ 91

RESULT 12  
Q6Y260 PRELIMINARY; PRT; 113 AA.  
ID Q6Y260; AC Q6Y260; DT 05-JUL-2004 (TREMBlrel. 27, Created)  
DT 05-JUL-2004 (TREMBlrel. 27, Last sequence update)  
DT 05-JUL-2004 (TREMBlrel. 27, Last annotation update)  
DE Alpha-actin protein (Fragment).  
OS Pagrus major (Red sea bream) (Chrysophrys major).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;  
OC Acanthomorpha; Acanthopterygii; Percomorpha; Perciformes; Percoidae;  
OC Sparidae; Pagrus.  
OX NCBI\_TaxID=143350;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC TISSUE=Spleen;  
RA Chen S.L., Xu M.Y., Hu S.L., Li L.;  
RT "Analysis of immune-relevant genes expressed in red sea bream spleen.";  
RL Aquaculture 240:115-130 (2004).  
DR EMBL; AY190676; AAF20152.1; -.  
DR GO:0005884; C:actin filament; IEA.  
DR GO:0003774; F:motor activity; IEA.  
DR GO:0005200; F:structural constituent of cytoskeleton; IEA.  
DR InterPro; IPR004001; Actin.  
DR PROSITE; PS00432; ACTINS\_2; 1.  
FT NON\_TER 1  
SQ SEQUENCE 113 AA; 12550 MW; 4C796A5B5B35269F CRC64;

Query Match 100.0%; Score 75; DB 2; Length 113;  
Best Local Similarity 100.0%; Pred. No. 2.9e-05; Indels 0; Gaps 0;  
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 LASLSTFQQMWSKQ 15  
Db 84 LASLSTFQQMWSKQ 98

RESULT 13  
Q6UIS1 PRELIMINARY; PRT; 123 AA.  
ID Q6UIS1; AC Q6UIS1; DT 05-JUL-2004 (TREMBlrel. 27, Created)  
DT 05-JUL-2004 (TREMBlrel. 27, Last sequence update)  
DT 05-JUL-2004 (TREMBlrel. 27, Last annotation update)  
DE Actin beta subunit (Fragment).  
GN Name=ACTB;



OS Macaca mulatta (Rhesus macaque).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Primates; Catarrhini; Cercopitheciidae;  
 OC Cercopitheciinae; Macaca.  
 OX NCBI\_TaxID=9544;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=22946167; PubMed=14557539; DOI=10.1073/pnas.2135499100;  
 RA Caceres M., Lachuer J., Zapala M.A., Redmond J.C., Kudo L.,  
 RA Geschwind D.H., Lockhart D.J., Preuss T.M., Barlow C.;  
 RT "Elevated gene expression levels distinguish human from non-human  
 RT primate brains.";  
 RL Proc. Natl. Acad. Sci. U.S.A. 100:13030-13035(2003).  
 DR EMBL; AY369786; AAR1122.1; -;  
 DR GO; GO:0005884; C:actin filament; IEA.  
 DR GO; GO:0003774; F:motor activity; IEA.  
 DR GO; GO:0005200; F:structural constituent of cytoskeleton; IEA.  
 DR InterPro; IPR004001; Actin.  
 DR InterPro; IPR004000; Actin\_like.  
 DR SMART; SM00268; ACTIN; 1.  
 DR PROSITE; PS00432; ACTINS\_2; 1.  
 FT NON TER 1  
 SQ SEQUENCE 123 AA; 13874 MW; BBOA1DA9C05599E1 CRC64;

Query Match 100.0%; Score 75; DB 2; Length 123;  
 Best Local Similarity 100.0%; Pred. No. 3.2e-05;  
 Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 LASLSTFOQMWSKQ 15  
 |||||  
 Db 94 LASLSTFOQMWSKQ 108

RESULT 14  
 Q6UIS2 PRELIMINARY; PRT; 123 AA.  
 AC Q6UIS2;  
 DT 05-JUL-2004 (TReMBLrel. 27, Created)  
 DT 05-JUL-2004 (TReMBLrel. 27, Last sequence update)  
 DT 05-JUL-2004 (TReMBLrel. 27, Last annotation update)  
 DE Actin beta subunit (Fragment).  
 GN Name=ACTB;  
 OS Pan troglodytes (Chimpanzee).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Pan.  
 OX NCBI\_TaxID=9598;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=22946167; PubMed=14557539; DOI=10.1073/pnas.2135499100;  
 RA Caceres M., Lachuer J., Zapala M.A., Redmond J.C., Kudo L.,  
 RA Geschwind D.H., Lockhart D.J., Preuss T.M., Barlow C.;  
 RT "Elevated gene expression levels distinguish human from non-human  
 RT primate brains.";  
 RL Proc. Natl. Acad. Sci. U.S.A. 100:13030-13035(2003).  
 DR EMBL; AY369785; AAR1122.1; -;  
 DR GO; GO:0005884; C:actin filament; IEA.  
 DR GO; GO:0003774; F:motor activity; IEA.  
 DR GO; GO:0005200; F:structural constituent of cytoskeleton; IEA.  
 DR InterPro; IPR004001; Actin.  
 DR InterPro; IPR004000; Actin\_like.  
 DR SMART; SM00268; ACTIN; 1.  
 DR PROSITE; PS00432; ACTINS\_2; 1.  
 FT NON TER 1  
 SQ SEQUENCE 123 AA; 13874 MW; BBOA1DA9C05599E1 CRC64;

Query Match 100.0%; Score 75; DB 2; Length 123;  
 Best Local Similarity 100.0%; Pred. No. 3.2e-05;  
 Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 LASLSTFOQMWSKQ 15  
 |||||  
 Db 94 LASLSTFOQMWSKQ 108

RESULT 15  
 Q9DFK3 PRELIMINARY; PRT; 124 AA.  
 AC Q9DFK3;  
 DT 01-MAR-2001 (TReMBLrel. 16, Created)  
 DT 01-MAR-2001 (TReMBLrel. 16, Last sequence update)  
 DT 01-MAR-2004 (TReMBLrel. 26, Last annotation update)  
 DE Skeletal alpha-actin (Fragment).  
 OS Gillichthys mirabilis (Long-jawed mudsucker).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;  
 OC Acanthomorpha; Acanthopterygii; Perciformes; Gobioidae;  
 OC Gobiidae; Gillichthys.  
 OX NCBI\_TaxID=8222;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=Skeletal white muscle;  
 RX MEDLINE=21117151; PubMed=11172064; DOI=10.1073/pnas.98.4.1993;  
 RA Gracey A.Y., Troll J.V., Somero G.N.;  
 RT "Hypoxia-induced gene expression profiling in the euryoxic fish  
 RT Gillichthys mirabilis.";  
 RL Proc. Natl. Acad. Sci. U.S.A. 98:1993-1998(2001).  
 DR EMBL; AF286232; AAG13352.1; -;  
 DR HSSP; P02568; 1025.  
 DR GO; GO:0015629; C:actin cytoskeleton; IEA.  
 DR GO; GO:0005200; F:structural constituent of cytoskeleton; IEA.  
 DR InterPro; IPR004000; Actin\_like.  
 DR Pfam; PF00022; Actin; 1.  
 FT NON TER 1  
 SQ SEQUENCE 124 AA; 13964 MW; 6BA11E35FD5DC323 CRC64;

Query Match 100.0%; Score 75; DB 2; Length 124;  
 Best Local Similarity 100.0%; Pred. No. 3.2e-05;  
 Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 LASLSTFOQMWSKQ 15  
 |||||  
 Db 95 LASLSTFOQMWSKQ 109

Search completed: April 8, 2005, 12:03:11  
 Job time : 58.8667 secs

**THIS PAGE BLANK (USPTO)**

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: April 8, 2005, 09:58:56 ; Search time 62.4 seconds  
(without alignments)

92.971 Million cell updates/sec

Title: US-09-423-351C-9

Perfect score: 76

Sequence: 1 DEAQSKRGILTLYKYP 15

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 2105692 seqs, 386760381 residues

Total number of hits satisfying chosen parameters: 2105692

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : A Geneseq\_16Dec04:\*

- 1: Geneseqp1980s:\*
- 2: Geneseqp1990s:\*
- 3: Geneseqp2000s:\*
- 4: Geneseqp2001s:\*
- 5: Geneseqp2002s:\*
- 6: Geneseqp2003as:\*
- 7: Geneseqp2003bs:\*
- 8: Geneseqp2004s:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	76	100.0	15	2	AAW92535 Beta-acti
2	76	100.0	54	3	AAG18905
3	76	100.0	58	3	AAG18945
4	76	100.0	73	3	AAG27057
5	76	100.0	74	3	AAG18936
6	76	100.0	75	3	AAG18944
7	76	100.0	79	3	AAG12594
8	76	100.0	82	3	AAG18904
9	76	100.0	88	3	AAG18916
10	76	100.0	91	3	AAG18935
11	76	100.0	95	3	AAG19142
12	76	100.0	96	3	AAG12593
13	76	100.0	99	3	AAG18903
14	76	100.0	101	3	AAG27056
15	76	100.0	105	5	ABP42469
16	76	100.0	108	4	AAm18324
17	76	100.0	108	4	ABP37358
18	76	100.0	108	4	AAm30812
19	76	100.0	108	4	ABP32106
20	76	100.0	108	4	ABP22644
21	76	100.0	108	4	AAm70488
22	76	100.0	108	4	AAm58048
23	76	100.0	108	4	ABG52170
24	76	100.0	108	4	AAm05932
25	76	100.0	108	5	ABG40127

26	76	100.0	111	3	AAG19141
27	76	100.0	116	3	AAG18915
28	76	100.0	118	3	AAG27055
29	76	100.0	133	3	AAG18914
30	76	100.0	150	4	ABG15101
31	76	100.0	157	4	ABP66853
32	76	100.0	166	4	ABG26869
33	76	100.0	167	4	AAU33077
34	76	100.0	201	5	ABP43176
35	76	100.0	204	5	ABP43016
36	76	100.0	211	7	ABM73899
37	76	100.0	219	8	ABO58422
38	76	100.0	239	7	ABE31206
39	76	100.0	258	4	ABG00383
40	76	100.0	276	3	AAG50947
41	76	100.0	277	3	AAG05115
42	76	100.0	304	3	AAG50946
43	76	100.0	305	3	AAG05114
44	76	100.0	308	3	AAG37210
45	76	100.0	311	3	AAG37209

#### ALIGNMENTS

RESULT 1

AAW92535

ID AAW92535 standard; peptide; 15 AA.

AC AAW92535;

DT 26-APR-1999 (first entry)

DE Beta-actin reference peptide substrate #9.

XX Peptide substrate; CCT; eukaryotic type II chaperonin complex; cyclin;  
KW binding agent; substrate-binding site; SBS; substrate folding; actin;  
KW tubulin; treatment; cancer; anticancer drug; viral infection; screening;  
KW reduced toxicity.

XX Synthetic.

XX WO9853322-A1.

XX 26-NOV-1998.

XX 22-MAY-1998; 98WO-GB001485.

XX 23-MAY-1997; 97GB-00010762.

XX (CANC-) INST CANCER RES ROYAL CANCER HOSPITAL.

XX Willison K, Hynes G, Liou AK;

XX WPI; 1999-070162/06.

XX Identifying specific binding agents for substrate binding site in CCT  
PT chaperonin complex - also new peptide binding agents and their mimetics,  
PT and peptides containing a specific CCT binding site, used for treating  
cancer.

XX Disclosure; Fig 10; 97pp; English.

XX This invention describes a method which uses the CCT (eukaryotic type II  
CC chaperonin) complex or part of it, for identifying a binding agent that  
CC can occupy a substrate-binding site (SBS) on the CCT complex. By binding  
CC to the CCT complex, the binding agents block an SBS so that biological  
CC activity of the CCT complex is affected, particularly its ability to fold  
CC substrates such as actin, tubulin and cyclin. The binding agents are  
CC useful for treatment of cancer, particularly when used in combination  
CC with an anticancer drug, or viral infections. Nucleic acid fragments are  
CC used to screen for agents, e.g. binding agents that modulate interaction  
CC between the CCT complex and a protein that is to be folded. The binding

CC agents may target cells that are actively synthesising tubulin etc.  
 CC (unlike known microtubule-stabilising agents that affect all cells), so  
 CC should have reduced toxicity for normal cells. AAW92527-W92541 are  
 CC peptide substrates used in the method of the invention

XX Sequence 15 AA;

Query Match 100.0%; Score 76; DB 2; Length 15;  
 Best Local Similarity 100.0%; Pred. No. 7.9e-07;  
 Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 DEAQSKRGILTLKYP 15  
 |||||  
 Db 1 DEAQSKRGILTLKYP 15

# RESULT 2

AAG18905  
 ID AAG18905 standard; protein; 54 AA.

XX AC AAG18905;

XX DT 17-OCT-2000 (first entry)

XX DE Zea mays protein fragment SEQ ID NO: 20499.

XX KW Protein identification; signal transduction pathway; metabolic pathway;  
 KW hybridisation assay; genetic mapping; gene expression control; promoter;  
 KW termination sequence; corn.

XX OS Zea mays subsp. mays.

XX PN EP1033405-A2.

XX PD 06-SEP-2000.

XX PF 25-FEB-2000; 2000EP-00301439.

XX PR 25-FEB-1999; 99US-0121825P.  
 PR 05-MAR-1999; 99US-0123180P.  
 PR 09-MAR-1999; 99US-0123548P.  
 PR 23-MAR-1999; 99US-0125788P.  
 PR 25-MAR-1999; 99US-0126264P.  
 PR 29-MAR-1999; 99US-0126785P.  
 PR 01-APR-1999; 99US-0127462P.  
 PR 06-APR-1999; 99US-0128234P.  
 PR 08-APR-1999; 99US-0128714P.  
 PR 16-APR-1999; 99US-0129845P.  
 PR 19-APR-1999; 99US-0130077P.  
 PR 21-APR-1999; 99US-0130449P.  
 PR 23-APR-1999; 99US-0130510P.  
 PR 28-APR-1999; 99US-0130891P.  
 PR 30-APR-1999; 99US-0131449P.  
 PR 30-APR-1999; 99US-0132048P.  
 PR 30-APR-1999; 99US-0132407P.  
 PR 04-MAY-1999; 99US-0132484P.  
 PR 05-MAY-1999; 99US-0132485P.  
 PR 06-MAY-1999; 99US-0132486P.  
 PR 07-MAY-1999; 99US-0132487P.  
 PR 07-MAY-1999; 99US-0132863P.  
 PR 11-MAY-1999; 99US-0134256P.  
 PR 14-MAY-1999; 99US-0134218P.  
 PR 14-MAY-1999; 99US-0134219P.  
 PR 14-MAY-1999; 99US-0134221P.  
 PR 14-MAY-1999; 99US-0134370P.  
 PR 18-MAY-1999; 99US-0134768P.  
 PR 19-MAY-1999; 99US-0134941P.  
 PR 20-MAY-1999; 99US-0135124P.  
 PR 21-MAY-1999; 99US-0135353P.  
 PR 24-MAY-1999; 99US-0135629P.  
 PR 25-MAY-1999; 99US-0136021P.  
 PR 27-MAY-1999; 99US-0136392P.  
 PR 28-MAY-1999; 99US-0136782P.

PR 01-JUN-1999; 99US-0137222P.  
 PR 03-JUN-1999; 99US-0137528P.  
 PR 04-JUN-1999; 99US-0137502P.  
 PR 07-JUN-1999; 99US-0137724P.  
 PR 08-JUN-1999; 99US-0138094P.  
 PR 10-JUN-1999; 99US-0138540P.  
 PR 10-JUN-1999; 99US-0138847P.  
 PR 14-JUN-1999; 99US-0139119P.  
 PR 16-JUN-1999; 99US-0139452P.  
 PR 17-JUN-1999; 99US-0139453P.  
 PR 17-JUN-1999; 99US-0139492P.  
 PR 18-JUN-1999; 99US-0139454P.  
 PR 18-JUN-1999; 99US-0139455P.  
 PR 18-JUN-1999; 99US-0139456P.  
 PR 18-JUN-1999; 99US-0139457P.  
 PR 18-JUN-1999; 99US-0139458P.  
 PR 18-JUN-1999; 99US-0139459P.  
 PR 18-JUN-1999; 99US-0139460P.  
 PR 18-JUN-1999; 99US-0139461P.  
 PR 18-JUN-1999; 99US-0139462P.  
 PR 18-JUN-1999; 99US-0139463P.  
 PR 18-JUN-1999; 99US-0139750P.  
 PR 18-JUN-1999; 99US-0139763P.  
 PR 21-JUN-1999; 99US-0139817P.  
 PR 22-JUN-1999; 99US-0139859P.  
 PR 23-JUN-1999; 99US-0140353P.  
 PR 23-JUN-1999; 99US-0140354P.  
 PR 24-JUN-1999; 99US-0140695P.  
 PR 28-JUN-1999; 99US-0140823P.  
 PR 29-JUN-1999; 99US-0140991P.  
 PR 30-JUN-1999; 99US-0141287P.  
 PR 01-JUL-1999; 99US-0141842P.  
 PR 01-JUL-1999; 99US-0142154P.  
 PR 02-JUL-1999; 99US-0142055P.  
 PR 06-JUL-1999; 99US-0142390P.  
 PR 08-JUL-1999; 99US-0142803P.  
 PR 09-JUL-1999; 99US-0142920P.  
 PR 12-JUL-1999; 99US-0142977P.  
 PR 13-JUL-1999; 99US-0143542P.  
 PR 14-JUL-1999; 99US-0143624P.  
 PR 15-JUL-1999; 99US-0144005P.  
 PR 16-JUL-1999; 99US-0144085P.  
 PR 16-JUL-1999; 99US-0144086P.  
 PR 19-JUL-1999; 99US-0144325P.  
 PR 19-JUL-1999; 99US-0144331P.  
 PR 19-JUL-1999; 99US-0144332P.  
 PR 19-JUL-1999; 99US-0144333P.  
 PR 19-JUL-1999; 99US-0144334P.  
 PR 19-JUL-1999; 99US-0144335P.  
 PR 20-JUL-1999; 99US-0144352P.  
 PR 20-JUL-1999; 99US-0144632P.  
 PR 20-JUL-1999; 99US-0144884P.  
 PR 21-JUL-1999; 99US-0144814P.  
 PR 21-JUL-1999; 99US-0145086P.  
 PR 21-JUL-1999; 99US-0145088P.  
 PR 21-JUL-1999; 99US-0145085P.  
 PR 22-JUL-1999; 99US-0145087P.  
 PR 22-JUL-1999; 99US-0145089P.  
 PR 22-JUL-1999; 99US-0145192P.  
 PR 23-JUL-1999; 99US-0145145P.  
 PR 23-JUL-1999; 99US-0145218P.  
 PR 23-JUL-1999; 99US-0145224P.  
 PR 26-JUL-1999; 99US-0145276P.  
 PR 27-JUL-1999; 99US-0145913P.  
 PR 27-JUL-1999; 99US-0145918P.  
 PR 27-JUL-1999; 99US-0145919P.  
 PR 02-AUG-1999; 99US-0145951P.  
 PR 02-AUG-1999; 99US-0146386P.  
 PR 02-AUG-1999; 99US-0146388P.  
 PR 02-AUG-1999; 99US-0146389P.  
 PR 03-AUG-1999; 99US-0147038P.  
 PR 04-AUG-1999; 99US-0147204P.  
 PR 04-AUG-1999; 99US-0147302P.

05-AUG-1999; 99US-0147192P.  
05-AUG-1999; 99US-0147260P.  
06-AUG-1999; 99US-0147303P.  
06-AUG-1999; 99US-0147416P.  
09-AUG-1999; 99US-0147493P.  
09-AUG-1999; 99US-0147933P.  
10-AUG-1999; 99US-0148171P.  
11-AUG-1999; 99US-0148319P.  
12-AUG-1999; 99US-0148341P.  
13-AUG-1999; 99US-0148565P.  
13-AUG-1999; 99US-0148684P.  
16-AUG-1999; 99US-0149368P.  
17-AUG-1999; 99US-0149175P.  
18-AUG-1999; 99US-0149426P.  
20-AUG-1999; 99US-0149722P.  
20-AUG-1999; 99US-0149723P.  
20-AUG-1999; 99US-0149923P.  
23-AUG-1999; 99US-0149902P.  
23-AUG-1999; 99US-0149930P.  
25-AUG-1999; 99US-0150566P.  
26-AUG-1999; 99US-0150884P.  
27-AUG-1999; 99US-0151065P.  
27-AUG-1999; 99US-0151068P.  
27-AUG-1999; 99US-0151080P.  
30-AUG-1999; 99US-0151303P.  
31-AUG-1999; 99US-0151438P.  
01-SEP-1999; 99US-0151930P.  
07-SEP-1999; 99US-0152363P.  
10-SEP-1999; 99US-0153070P.  
13-SEP-1999; 99US-0153758P.  
15-SEP-1999; 99US-0154018P.  
16-SEP-1999; 99US-0154039P.  
20-SEP-1999; 99US-0154779P.  
22-SEP-1999; 99US-0155133P.  
23-SEP-1999; 99US-0155486P.  
24-SEP-1999; 99US-0155659P.  
28-SEP-1999; 99US-0156458P.  
29-SEP-1999; 99US-0156596P.  
04-OCT-1999; 99US-0157117P.  
05-OCT-1999; 99US-0157753P.  
06-OCT-1999; 99US-0157863P.  
07-OCT-1999; 99US-0158029P.  
08-OCT-1999; 99US-0158232P.  
12-OCT-1999; 99US-0159369P.  
13-OCT-1999; 99US-0159293P.  
13-OCT-1999; 99US-0159294P.  
13-OCT-1999; 99US-0159295P.  
14-OCT-1999; 99US-0159329P.  
14-OCT-1999; 99US-0159330P.  
14-OCT-1999; 99US-0159331P.  
14-OCT-1999; 99US-0159637P.  
14-OCT-1999; 99US-0159638P.  
18-OCT-1999; 99US-0159584P.  
21-OCT-1999; 99US-0160741P.  
21-OCT-1999; 99US-0160767P.  
21-OCT-1999; 99US-0160768P.  
21-OCT-1999; 99US-0160770P.  
21-OCT-1999; 99US-0160814P.  
21-OCT-1999; 99US-0160815P.  
22-OCT-1999; 99US-0160980P.  
22-OCT-1999; 99US-0160981P.  
22-OCT-1999; 99US-0160989P.  
25-OCT-1999; 99US-0161404P.  
25-OCT-1999; 99US-0161405P.  
26-OCT-1999; 99US-0161406P.  
26-OCT-1999; 99US-0161359P.  
26-OCT-1999; 99US-0161360P.  
26-OCT-1999; 99US-0161361P.  
28-OCT-1999; 99US-0161920P.  
28-OCT-1999; 99US-0161992P.  
28-OCT-1999; 99US-0161993P.  
29-OCT-1999; 99US-0162142P.  
29-OCT-1999; 99US-0162143P.

Query Match 100.0%; Score 76; DB 3; Length 54;  
Best Local Similarity 100.0%; Pred. No. 3.2e-06;  
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 DEAQSKRGILTLYP 15  
| | | | | | | | | | | | | | |  
Db 13 DEAQSKRGILTLYP 27

RESULT 3  
AAG18945  
ID AAG18945 standard; protein; 58 AA.  
XX  
AC AAG18945;  
XX  
DT 17-OCT-2000 (first entry)  
XX  
DE Zea mays protein fragment SEQ ID NO: 20555.  
XX  
KW Protein identification; signal transduction pathway; metabolic pathway;  
KW hybridisation assay; genetic mapping; gene expression control; promoter;  
KW termination sequence; corn.  
XX  
OS Zea mays subsp. mays.  
XX  
PN EP1033405-A2.  
XX  
PD 06-SEP-2000.  
XX  
PF 25-FEB-2000; 2000EP-00301439.  
XX  
PR 25-FEB-1999; 99US-0121825P.  
PR 05-MAR-1999; 99US-0123180P.  
PR 09-MAR-1999; 99US-0123548P.  
PR 23-MAR-1999; 99US-0125788P.  
PR 25-MAR-1999; 99US-0126264P.  
PR 29-MAR-1999; 99US-0126785P.  
PR 01-APR-1999; 99US-0127462P.  
PR 06-APR-1999; 99US-0128234P.  
PR 08-APR-1999; 99US-0128714P.  
PR 16-APR-1999; 99US-0129845P.  
PR 19-APR-1999; 99US-0130077P.  
PR 21-APR-1999; 99US-0130449P.  
PR 23-APR-1999; 99US-0130510P.  
PR 23-APR-1999; 99US-0130891P.  
PR 28-APR-1999; 99US-0131443P.  
PR 30-APR-1999; 99US-0132048P.  
PR 30-APR-1999; 99US-0132407P.  
PR 04-MAY-1999; 99US-0132484P.  
PR 05-MAY-1999; 99US-0132485P.  
PR 06-MAY-1999; 99US-0132486P.  
PR 06-MAY-1999; 99US-0132487P.  
PR 07-MAY-1999; 99US-0132863P.  
PR 11-MAY-1999; 99US-0134256P.  
PR 14-MAY-1999; 99US-0134218P.  
PR 14-MAY-1999; 99US-0134219P.  
PR 14-MAY-1999; 99US-0134221P.  
PR 14-MAY-1999; 99US-0134370P.  
PR 18-MAY-1999; 99US-0134768P.  
PR 19-MAY-1999; 99US-0134941P.  
PR 20-MAY-1999; 99US-0135124P.  
PR 21-MAY-1999; 99US-0135353P.  
PR 24-MAY-1999; 99US-0135629P.  
PR 25-MAY-1999; 99US-0136021P.  
PR 27-MAY-1999; 99US-0136392P.  
PR 28-MAY-1999; 99US-0136782P.  
PR 01-JUN-1999; 99US-0137222P.  
PR 03-JUN-1999; 99US-0137528P.  
PR 04-JUN-1999; 99US-0137502P.  
PR 07-JUN-1999; 99US-0137724P.  
PR 08-JUN-1999; 99US-0138094P.  
PR 10-JUN-1999; 99US-0138540P.  
PR 10-JUN-1999; 99US-0138847P.

PR 14-JUN-1999; 99US-0139119P.  
 PR 16-JUN-1999; 99US-0139452P.  
 PR 16-JUN-1999; 99US-0139453P.  
 PR 17-JUN-1999; 99US-0139492P.  
 PR 18-JUN-1999; 99US-0139454P.  
 PR 18-JUN-1999; 99US-0139455P.  
 PR 18-JUN-1999; 99US-0139456P.  
 PR 18-JUN-1999; 99US-0139457P.  
 PR 18-JUN-1999; 99US-0139458P.  
 PR 18-JUN-1999; 99US-0139459P.  
 PR 18-JUN-1999; 99US-0139460P.  
 PR 18-JUN-1999; 99US-0139461P.  
 PR 18-JUN-1999; 99US-0139462P.  
 PR 18-JUN-1999; 99US-0139463P.  
 PR 18-JUN-1999; 99US-0139750P.  
 PR 18-JUN-1999; 99US-0139763P.  
 PR 21-JUN-1999; 99US-0139817P.  
 PR 22-JUN-1999; 99US-0139899P.  
 PR 23-JUN-1999; 99US-0140353P.  
 PR 23-JUN-1999; 99US-0140354P.  
 PR 24-JUN-1999; 99US-0140695P.  
 PR 28-JUN-1999; 99US-0140823P.  
 PR 29-JUN-1999; 99US-0140991P.  
 PR 30-JUN-1999; 99US-0141287P.  
 PR 01-JUL-1999; 99US-0141842P.  
 PR 01-JUL-1999; 99US-0142154P.  
 PR 02-JUL-1999; 99US-0142055P.  
 PR 06-JUL-1999; 99US-0142390P.  
 PR 08-JUL-1999; 99US-0142803P.  
 PR 09-JUL-1999; 99US-0142920P.  
 PR 12-JUL-1999; 99US-0142977P.  
 PR 13-JUL-1999; 99US-0143543P.  
 PR 14-JUL-1999; 99US-0143624P.  
 PR 15-JUL-1999; 99US-0144005P.  
 PR 16-JUL-1999; 99US-0144085P.  
 PR 16-JUL-1999; 99US-0144086P.  
 PR 19-JUL-1999; 99US-0144325P.  
 PR 19-JUL-1999; 99US-0144331P.  
 PR 19-JUL-1999; 99US-0144332P.  
 PR 19-JUL-1999; 99US-0144333P.  
 PR 19-JUL-1999; 99US-0144334P.  
 PR 19-JUL-1999; 99US-0144335P.  
 PR 20-JUL-1999; 99US-0144332P.  
 PR 20-JUL-1999; 99US-0144363P.  
 PR 20-JUL-1999; 99US-0144884P.  
 PR 21-JUL-1999; 99US-0144814P.  
 PR 21-JUL-1999; 99US-0145086P.  
 PR 21-JUL-1999; 99US-0145088P.  
 PR 22-JUL-1999; 99US-0145087P.  
 PR 22-JUL-1999; 99US-0145089P.  
 PR 22-JUL-1999; 99US-0145192P.  
 PR 23-JUL-1999; 99US-0145115P.  
 PR 23-JUL-1999; 99US-0145218P.  
 PR 23-JUL-1999; 99US-0145224P.  
 PR 23-JUL-1999; 99US-0145224P.  
 PR 26-JUL-1999; 99US-0145227P.  
 PR 27-JUL-1999; 99US-0145913P.  
 PR 27-JUL-1999; 99US-0145918P.  
 PR 27-JUL-1999; 99US-0145919P.  
 PR 28-JUL-1999; 99US-0145951P.  
 PR 02-AUG-1999; 99US-0146386P.  
 PR 02-AUG-1999; 99US-0146388P.  
 PR 02-AUG-1999; 99US-0146389P.  
 PR 03-AUG-1999; 99US-0147038P.  
 PR 04-AUG-1999; 99US-0147204P.  
 PR 04-AUG-1999; 99US-0147302P.  
 PR 05-AUG-1999; 99US-0147192P.  
 PR 05-AUG-1999; 99US-0147260P.  
 PR 06-AUG-1999; 99US-0147303P.  
 PR 06-AUG-1999; 99US-0147416P.  
 PR 09-AUG-1999; 99US-0147493P.  
 PR 09-AUG-1999; 99US-0147935P.  
 PR 10-AUG-1999; 99US-0148171P.

PR 11-AUG-1999; 99US-0148319P.  
 PR 12-AUG-1999; 99US-0148341P.  
 PR 13-AUG-1999; 99US-0148565P.  
 PR 13-AUG-1999; 99US-0148684P.  
 PR 16-AUG-1999; 99US-0149368P.  
 PR 17-AUG-1999; 99US-0149175P.  
 PR 18-AUG-1999; 99US-0149426P.  
 PR 20-AUG-1999; 99US-0149722P.  
 PR 20-AUG-1999; 99US-0149723P.  
 PR 20-AUG-1999; 99US-0149929P.  
 PR 23-AUG-1999; 99US-0149902P.  
 PR 23-AUG-1999; 99US-0149930P.  
 PR 25-AUG-1999; 99US-0150566P.  
 PR 26-AUG-1999; 99US-0150884P.  
 PR 27-AUG-1999; 99US-0151065P.  
 PR 27-AUG-1999; 99US-0151066P.  
 PR 27-AUG-1999; 99US-0151080P.  
 PR 30-AUG-1999; 99US-0151303P.  
 PR 31-AUG-1999; 99US-0151438P.  
 PR 01-SEP-1999; 99US-0151930P.  
 PR 07-SEP-1999; 99US-0152363P.  
 PR 10-SEP-1999; 99US-0153070P.  
 PR 13-SEP-1999; 99US-0153758P.  
 PR 15-SEP-1999; 99US-0154018P.  
 PR 16-SEP-1999; 99US-0154039P.  
 PR 20-SEP-1999; 99US-0154779P.  
 PR 22-SEP-1999; 99US-0155139P.  
 PR 23-SEP-1999; 99US-0155486P.  
 PR 24-SEP-1999; 99US-0155659P.  
 PR 28-SEP-1999; 99US-0156458P.  
 PR 29-SEP-1999; 99US-0156596P.  
 PR 04-OCT-1999; 99US-0157117P.  
 PR 05-OCT-1999; 99US-0157753P.  
 PR 06-OCT-1999; 99US-0157865P.  
 PR 07-OCT-1999; 99US-0158029P.  
 PR 08-OCT-1999; 99US-0158232P.  
 PR 12-OCT-1999; 99US-0158369P.  
 PR 13-OCT-1999; 99US-0159293P.  
 PR 13-OCT-1999; 99US-0159294P.  
 PR 13-OCT-1999; 99US-0159295P.  
 PR 14-OCT-1999; 99US-0159329P.  
 PR 14-OCT-1999; 99US-0159330P.  
 PR 14-OCT-1999; 99US-0159331P.  
 PR 14-OCT-1999; 99US-0159637P.  
 PR 14-OCT-1999; 99US-0159638P.  
 PR 18-OCT-1999; 99US-0159584P.  
 PR 21-OCT-1999; 99US-0160741P.  
 PR 21-OCT-1999; 99US-0160767P.  
 PR 21-OCT-1999; 99US-0160768P.  
 PR 21-OCT-1999; 99US-0160770P.  
 PR 21-OCT-1999; 99US-0160814P.  
 PR 21-OCT-1999; 99US-0160815P.  
 PR 22-OCT-1999; 99US-0160980P.  
 PR 22-OCT-1999; 99US-0160981P.  
 PR 22-OCT-1999; 99US-0160989P.  
 PR 25-OCT-1999; 99US-0161404P.  
 PR 25-OCT-1999; 99US-0161405P.  
 PR 25-OCT-1999; 99US-0161406P.  
 PR 26-OCT-1999; 99US-0161359P.  
 PR 26-OCT-1999; 99US-0161360P.  
 PR 26-OCT-1999; 99US-0161361P.  
 PR 28-OCT-1999; 99US-0161920P.  
 PR 28-OCT-1999; 99US-0161920P.  
 PR 28-OCT-1999; 99US-0161920P.  
 PR 28-OCT-1999; 99US-0161933P.  
 PR 29-OCT-1999; 99US-0162142P.

Query Match 100.0%; Score 76; DB 3; Length 58;  
 Best Local Similarity 100.0%; Pred. No. 3.4e-06;  
 Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 DEAQSKRGILTLKYP 15  
 |||||  
 Db 41 DEAQSKRGILTLKYP 55

RESULT 4  
AAG27057  
XX ID AAG27057 standard; protein; 73 AA.  
XX AC AAG27057;  
XX DT 17-OCT-2000 (first entry)  
XX DE Zea mays protein fragment SEQ ID NO: 31747.  
XX KW Protein identification; signal transduction pathway; metabolic pathway;  
XX KW hybridisation assay; genetic mapping; gene expression control; promoter;  
XX KW termination sequence; corn.  
XX OS Zea mays subsp. mays.  
XX PN EP1033405-A2.  
XX PD 06-SEP-2000.  
XX PF 25-FEB-2000; 2000EP-00301439.  
XX PR 25-FEB-1999; 99US-0121825P.  
PR 05-MAR-1999; 99US-0123180P.  
PR 09-MAR-1999; 99US-0123548P.  
PR 23-MAR-1999; 99US-0125788P.  
PR 25-MAR-1999; 99US-0126264P.  
PR 29-MAR-1999; 99US-0126783P.  
PR 01-APR-1999; 99US-0127462P.  
PR 06-APR-1999; 99US-0128234P.  
PR 08-APR-1999; 99US-0128714P.  
PR 16-APR-1999; 99US-0129845P.  
PR 19-APR-1999; 99US-0130077P.  
PR 21-APR-1999; 99US-0130449P.  
PR 23-APR-1999; 99US-0130510P.  
PR 28-APR-1999; 99US-0130891P.  
PR 28-APR-1999; 99US-0131449P.  
PR 30-APR-1999; 99US-0132048P.  
PR 30-APR-1999; 99US-0132407P.  
PR 04-MAY-1999; 99US-0132484P.  
PR 05-MAY-1999; 99US-0132485P.  
PR 06-MAY-1999; 99US-0132486P.  
PR 06-MAY-1999; 99US-0132487P.  
PR 07-MAY-1999; 99US-0132863P.  
PR 11-MAY-1999; 99US-0134256P.  
PR 14-MAY-1999; 99US-0134218P.  
PR 14-MAY-1999; 99US-0134219P.  
PR 14-MAY-1999; 99US-0134221P.  
PR 14-MAY-1999; 99US-0134370P.  
PR 18-MAY-1999; 99US-0134768P.  
PR 19-MAY-1999; 99US-0134941P.  
PR 20-MAY-1999; 99US-0135124P.  
PR 21-MAY-1999; 99US-0135353P.  
PR 24-MAY-1999; 99US-0135629P.  
PR 25-MAY-1999; 99US-0136021P.  
PR 27-MAY-1999; 99US-0136392P.  
PR 28-MAY-1999; 99US-0136782P.  
PR 01-JUN-1999; 99US-0137222P.  
PR 03-JUN-1999; 99US-0137528P.  
PR 04-JUN-1999; 99US-0137502P.  
PR 07-JUN-1999; 99US-0137724P.  
PR 08-JUN-1999; 99US-0138094P.  
PR 10-JUN-1999; 99US-0138540P.  
PR 10-JUN-1999; 99US-0138847P.  
PR 14-JUN-1999; 99US-0139119P.  
PR 16-JUN-1999; 99US-0139452P.  
PR 16-JUN-1999; 99US-0139453P.  
PR 17-JUN-1999; 99US-0139492P.  
PR 18-JUN-1999; 99US-0139454P.  
PR 18-JUN-1999; 99US-0139455P.  
PR 18-JUN-1999; 99US-0139456P.  
PR 18-JUN-1999; 99US-0139457P.  
PR 18-JUN-1999; 99US-0139458P.  
PR 18-JUN-1999; 99US-0139459P.  
PR 18-JUN-1999; 99US-0139460P.  
PR 18-JUN-1999; 99US-0139461P.  
PR 18-JUN-1999; 99US-0139462P.  
PR 18-JUN-1999; 99US-0139463P.  
PR 18-JUN-1999; 99US-0139750P.  
PR 18-JUN-1999; 99US-0139763P.  
PR 21-JUN-1999; 99US-0139817P.  
PR 22-JUN-1999; 99US-0139899P.  
PR 23-JUN-1999; 99US-0140353P.  
PR 23-JUN-1999; 99US-0140354P.  
PR 24-JUN-1999; 99US-0140695P.  
PR 28-JUN-1999; 99US-0140823P.  
PR 29-JUN-1999; 99US-0140991P.  
PR 30-JUN-1999; 99US-0141287P.  
PR 01-JUL-1999; 99US-0141842P.  
PR 01-JUL-1999; 99US-0142154P.  
PR 02-JUL-1999; 99US-0142055P.  
PR 06-JUL-1999; 99US-0142390P.  
PR 08-JUL-1999; 99US-0142803P.  
PR 09-JUL-1999; 99US-0142920P.  
PR 12-JUL-1999; 99US-0142977P.  
PR 13-JUL-1999; 99US-0143542P.  
PR 14-JUL-1999; 99US-0143624P.  
PR 15-JUL-1999; 99US-0144005P.  
PR 16-JUL-1999; 99US-0144085P.  
PR 16-JUL-1999; 99US-0144086P.  
PR 19-JUL-1999; 99US-0144325P.  
PR 19-JUL-1999; 99US-0144331P.  
PR 19-JUL-1999; 99US-0144332P.  
PR 19-JUL-1999; 99US-0144333P.  
PR 19-JUL-1999; 99US-0144334P.  
PR 19-JUL-1999; 99US-0144335P.  
PR 20-JUL-1999; 99US-0144352P.  
PR 20-JUL-1999; 99US-0144632P.  
PR 21-JUL-1999; 99US-0144884P.  
PR 21-JUL-1999; 99US-0144814P.  
PR 21-JUL-1999; 99US-0145086P.  
PR 21-JUL-1999; 99US-0145088P.  
PR 22-JUL-1999; 99US-0145085P.  
PR 22-JUL-1999; 99US-0145087P.  
PR 22-JUL-1999; 99US-0145089P.  
PR 22-JUL-1999; 99US-0145192P.  
PR 23-JUL-1999; 99US-0145145P.  
PR 23-JUL-1999; 99US-0145218P.  
PR 23-JUL-1999; 99US-0145224P.  
PR 26-JUL-1999; 99US-0145276P.  
PR 27-JUL-1999; 99US-0145913P.  
PR 27-JUL-1999; 99US-0145918P.  
PR 27-JUL-1999; 99US-0145919P.  
PR 28-JUL-1999; 99US-0145951P.  
PR 02-AUG-1999; 99US-0146386P.  
PR 02-AUG-1999; 99US-0146388P.  
PR 02-AUG-1999; 99US-0146389P.  
PR 03-AUG-1999; 99US-0147038P.  
PR 04-AUG-1999; 99US-0147204P.  
PR 04-AUG-1999; 99US-0147302P.  
PR 05-AUG-1999; 99US-0147192P.  
PR 05-AUG-1999; 99US-0147260P.  
PR 06-AUG-1999; 99US-0147303P.  
PR 06-AUG-1999; 99US-0147416P.  
PR 09-AUG-1999; 99US-0147493P.  
PR 09-AUG-1999; 99US-0147935P.  
PR 10-AUG-1999; 99US-0148171P.  
PR 11-AUG-1999; 99US-0148319P.  
PR 12-AUG-1999; 99US-0148341P.  
PR 13-AUG-1999; 99US-0148565P.  
PR 13-AUG-1999; 99US-0148684P.  
PR 16-AUG-1999; 99US-0149368P.  
PR 17-AUG-1999; 99US-0149175P.  
PR 18-AUG-1999; 99US-0149426P.





```
PR 18-JUN-1999; 99US-0139750P.
PR 18-JUN-1999; 99US-0139763P.
PR 21-JUN-1999; 99US-0139817P.
PR 22-JUN-1999; 99US-0139899P.
PR 23-JUN-1999; 99US-0140353P.
PR 23-JUN-1999; 99US-0140354P.
PR 24-JUN-1999; 99US-0140695P.
PR 28-JUN-1999; 99US-0140823P.
PR 29-JUN-1999; 99US-0140991P.
PR 30-JUN-1999; 99US-0141287P.
PR 01-JUL-1999; 99US-0141842P.
PR 01-JUL-1999; 99US-0142154P.
PR 02-JUL-1999; 99US-0142055P.
PR 06-JUL-1999; 99US-0142390P.
PR 08-JUL-1999; 99US-0142803P.
PR 09-JUL-1999; 99US-0142920P.
PR 12-JUL-1999; 99US-0142977P.
PR 13-JUL-1999; 99US-0143542P.
PR 14-JUL-1999; 99US-0143624P.
PR 15-JUL-1999; 99US-0144005P.
PR 16-JUL-1999; 99US-0144085P.
PR 16-JUL-1999; 99US-0144086P.
PR 19-JUL-1999; 99US-0144325P.
PR 19-JUL-1999; 99US-0144331P.
PR 19-JUL-1999; 99US-0144332P.
PR 19-JUL-1999; 99US-0144333P.
PR 19-JUL-1999; 99US-0144333P.
PR 19-JUL-1999; 99US-0144334P.
PR 19-JUL-1999; 99US-0144335P.
PR 20-JUL-1999; 99US-0144352P.
PR 20-JUL-1999; 99US-0144632P.
PR 20-JUL-1999; 99US-0144884P.
PR 21-JUL-1999; 99US-0144814P.
PR 21-JUL-1999; 99US-0145086P.
PR 21-JUL-1999; 99US-0145088P.
PR 22-JUL-1999; 99US-0145085P.
PR 22-JUL-1999; 99US-0145087P.
PR 22-JUL-1999; 99US-0145089P.
PR 22-JUL-1999; 99US-0145192P.
PR 23-JUL-1999; 99US-0145154P.
PR 23-JUL-1999; 99US-0145218P.
PR 23-JUL-1999; 99US-0145219P.
PR 26-JUL-1999; 99US-0145224P.
PR 27-JUL-1999; 99US-0145276P.
PR 27-JUL-1999; 99US-0145913P.
PR 27-JUL-1999; 99US-0145918P.
PR 27-JUL-1999; 99US-0145919P.
PR 28-JUL-1999; 99US-0145951P.
PR 02-AUG-1999; 99US-0146386P.
PR 02-AUG-1999; 99US-0146388P.
PR 02-AUG-1999; 99US-0146389P.
PR 03-AUG-1999; 99US-0147030P.
PR 04-AUG-1999; 99US-0147204P.
PR 04-AUG-1999; 99US-0147302P.
PR 05-AUG-1999; 99US-0147192P.
PR 05-AUG-1999; 99US-0147260P.
PR 06-AUG-1999; 99US-0147303P.
PR 06-AUG-1999; 99US-0147416P.
PR 09-AUG-1999; 99US-0147493P.
PR 09-AUG-1999; 99US-0147935P.
PR 10-AUG-1999; 99US-0148171P.
PR 11-AUG-1999; 99US-0148319P.
PR 12-AUG-1999; 99US-0148341P.
PR 13-AUG-1999; 99US-0148565P.
PR 13-AUG-1999; 99US-0148684P.
PR 16-AUG-1999; 99US-0149368P.
PR 17-AUG-1999; 99US-0149175P.
PR 18-AUG-1999; 99US-0149426P.
PR 20-AUG-1999; 99US-0149722P.
PR 20-AUG-1999; 99US-0149723P.
PR 20-AUG-1999; 99US-0149929P.
PR 23-AUG-1999; 99US-0149902P.
PR 23-AUG-1999; 99US-0149930P.
PR 25-AUG-1999; 99US-0150566P.
PR 26-AUG-1999; 99US-0150884P.

PR 27-AUG-1999; 99US-0151065P.
PR 27-AUG-1999; 99US-0151066P.
PR 27-AUG-1999; 99US-0151080P.
PR 30-AUG-1999; 99US-0151303P.
PR 31-AUG-1999; 99US-0151438P.
PR 01-SEP-1999; 99US-0151930P.
PR 07-SEP-1999; 99US-0152363P.
PR 10-SEP-1999; 99US-0153070P.
PR 13-SEP-1999; 99US-0153758P.
PR 15-SEP-1999; 99US-0154018P.
PR 16-SEP-1999; 99US-0154039P.
PR 20-SEP-1999; 99US-0154779P.
PR 22-SEP-1999; 99US-0155139P.
PR 23-SEP-1999; 99US-0155486P.
PR 24-SEP-1999; 99US-0155659P.
PR 28-SEP-1999; 99US-0156458P.
PR 29-SEP-1999; 99US-0156596P.
PR 04-OCT-1999; 99US-0157117P.
PR 05-OCT-1999; 99US-015753P.
PR 06-OCT-1999; 99US-0157865P.
PR 07-OCT-1999; 99US-0158029P.
PR 08-OCT-1999; 99US-0158232P.
PR 12-OCT-1999; 99US-0158369P.
PR 13-OCT-1999; 99US-0159293P.
PR 13-OCT-1999; 99US-0159294P.
PR 13-OCT-1999; 99US-0159295P.
PR 14-OCT-1999; 99US-0159329P.
PR 14-OCT-1999; 99US-0159330P.
PR 14-OCT-1999; 99US-0159331P.
PR 14-OCT-1999; 99US-0159637P.
PR 14-OCT-1999; 99US-0159638P.
PR 18-OCT-1999; 99US-0159584P.
PR 21-OCT-1999; 99US-0160741P.
PR 21-OCT-1999; 99US-0160767P.
PR 21-OCT-1999; 99US-0160768P.
PR 21-OCT-1999; 99US-0160770P.
PR 21-OCT-1999; 99US-0160814P.
PR 21-OCT-1999; 99US-0160815P.
PR 22-OCT-1999; 99US-0160800P.
PR 22-OCT-1999; 99US-0160981P.
PR 22-OCT-1999; 99US-0160989P.
PR 25-OCT-1999; 99US-0161404P.
PR 25-OCT-1999; 99US-0161405P.
PR 25-OCT-1999; 99US-0161406P.
PR 26-OCT-1999; 99US-0161359P.
PR 26-OCT-1999; 99US-0161360P.
PR 26-OCT-1999; 99US-0161361P.
PR 28-OCT-1999; 99US-0161920P.
PR 28-OCT-1999; 99US-0161920P.
PR 28-OCT-1999; 99US-0161993P.
PR 29-OCT-1999; 99US-0162142P.

Query Match 100.0%; Score 76; DB 3; Length 74;
Best Local Similarity 100.0%; Pred. No. 4.5e-06;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 DEAQSKRGILTKYP 15
Db 41 DEAQSKRGILTKYP 55

RESULT 6
ID AAG18944 standard; protein; 75 AA.
XX AC AAG18944;
XX AC AAG18944;
DT 17-OCT-2000 (first entry)
XX DE Zea mays protein fragment SEQ ID NO: 20554.
XX KW Protein identification; signal transduction pathway; metabolic pathway;
KW hybridisation assay; genetic mapping; gene expression control; promoter;
```

KW	termination sequence; corn.	PR	28-JUN-1999;	99US-0140823P.
XX		PR	29-JUN-1999;	99US-0140991P.
OS		PR	30-JUN-1999;	99US-0141287P.
XX	Zea mays subsp. mays.	PR	01-JUL-1999;	99US-0141842P.
XX	EP1033405-A2.	PR	01-JUL-1999;	99US-0142154P.
XX		PR	02-JUL-1999;	99US-0142055P.
XX		PR	06-JUL-1999;	99US-0142390P.
XX		PR	08-JUL-1999;	99US-0142803P.
PF	06-SEP-2000.	PR	09-JUL-1999;	99US-0142920P.
XX		PR	12-JUL-1999;	99US-0142977P.
XX	25-FEB-2000; 2000EP-00301439.	PR	13-JUL-1999;	99US-0143542P.
XX		PR	14-JUL-1999;	99US-0143624P.
PR	25-FEB-1999;	PR	15-JUL-1999;	99US-0144005P.
PR	05-MAR-1999;	PR	16-JUL-1999;	99US-0144085P.
PR	09-MAR-1999;	PR	16-JUL-1999;	99US-0144086P.
PR	23-MAR-1999;	PR	16-JUL-1999;	99US-0144325P.
PR	25-MAR-1999;	PR	19-JUL-1999;	99US-0144331P.
PR	29-MAR-1999;	PR	19-JUL-1999;	99US-0144332P.
PR	01-APR-1999;	PR	19-JUL-1999;	99US-0144333P.
PR	06-APR-1999;	PR	19-JUL-1999;	99US-0144334P.
PR	08-APR-1999;	PR	19-JUL-1999;	99US-0144335P.
PR	16-APR-1999;	PR	19-JUL-1999;	99US-0144352P.
PR	19-APR-1999;	PR	20-JUL-1999;	99US-0144632P.
PR	21-APR-1999;	PR	20-JUL-1999;	99US-0144884P.
PR	23-APR-1999;	PR	21-JUL-1999;	99US-0144814P.
PR	28-APR-1999;	PR	21-JUL-1999;	99US-0145086P.
PR	30-APR-1999;	PR	21-JUL-1999;	99US-0145088P.
PR	30-APR-1999;	PR	22-JUL-1999;	99US-0145085P.
PR	04-MAY-1999;	PR	22-JUL-1999;	99US-0145087P.
PR	05-MAY-1999;	PR	22-JUL-1999;	99US-0145089P.
PR	06-MAY-1999;	PR	23-JUL-1999;	99US-0145192P.
PR	06-MAY-1999;	PR	23-JUL-1999;	99US-0145145P.
PR	07-MAY-1999;	PR	23-JUL-1999;	99US-0145218P.
PR	11-MAY-1999;	PR	23-JUL-1999;	99US-0145224P.
PR	14-MAY-1999;	PR	26-JUL-1999;	99US-0145276P.
PR	14-MAY-1999;	PR	27-JUL-1999;	99US-0145913P.
PR	14-MAY-1999;	PR	27-JUL-1999;	99US-0145918P.
PR	14-MAY-1999;	PR	27-JUL-1999;	99US-0145919P.
PR	18-MAY-1999;	PR	28-JUL-1999;	99US-0145951P.
PR	19-MAY-1999;	PR	28-JUL-1999;	99US-0146386P.
PR	20-MAY-1999;	PR	02-AUG-1999;	99US-0146388P.
PR	21-MAY-1999;	PR	02-AUG-1999;	99US-0146389P.
PR	21-MAY-1999;	PR	02-AUG-1999;	99US-0147038P.
PR	24-MAY-1999;	PR	03-AUG-1999;	99US-0147204P.
PR	25-MAY-1999;	PR	04-AUG-1999;	99US-0147302P.
PR	27-MAY-1999;	PR	04-AUG-1999;	99US-0147302P.
PR	28-MAY-1999;	PR	05-AUG-1999;	99US-0147192P.
PR	01-JUN-1999;	PR	05-AUG-1999;	99US-0147260P.
PR	03-JUN-1999;	PR	06-AUG-1999;	99US-0147303P.
PR	07-JUN-1999;	PR	06-AUG-1999;	99US-0147416P.
PR	08-JUN-1999;	PR	09-AUG-1999;	99US-0147935P.
PR	10-JUN-1999;	PR	09-AUG-1999;	99US-0147935P.
PR	10-JUN-1999;	PR	10-AUG-1999;	99US-0148171P.
PR	10-JUN-1999;	PR	11-AUG-1999;	99US-0148319P.
PR	14-JUN-1999;	PR	12-AUG-1999;	99US-0148341P.
PR	16-JUN-1999;	PR	13-AUG-1999;	99US-0148565P.
PR	16-JUN-1999;	PR	13-AUG-1999;	99US-0148684P.
PR	17-JUN-1999;	PR	16-AUG-1999;	99US-0149368P.
PR	18-JUN-1999;	PR	17-AUG-1999;	99US-0149175P.
PR	18-JUN-1999;	PR	18-AUG-1999;	99US-0149426P.
PR	18-JUN-1999;	PR	20-AUG-1999;	99US-0149722P.
PR	18-JUN-1999;	PR	20-AUG-1999;	99US-0149723P.
PR	18-JUN-1999;	PR	23-AUG-1999;	99US-0149929P.
PR	18-JUN-1999;	PR	23-AUG-1999;	99US-0149902P.
PR	18-JUN-1999;	PR	23-AUG-1999;	99US-0149910P.
PR	18-JUN-1999;	PR	25-AUG-1999;	99US-0150866P.
PR	18-JUN-1999;	PR	26-AUG-1999;	99US-0150884P.
PR	18-JUN-1999;	PR	27-AUG-1999;	99US-0151065P.
PR	18-JUN-1999;	PR	27-AUG-1999;	99US-0151066P.
PR	21-JUN-1999;	PR	27-AUG-1999;	99US-0151080P.
PR	21-JUN-1999;	PR	30-AUG-1999;	99US-0151303P.
PR	22-JUN-1999;	PR	31-AUG-1999;	99US-0151438P.
PR	23-JUN-1999;	PR	01-SEP-1999;	99US-0151930P.
PR	23-JUN-1999;	PR	07-SEP-1999;	99US-0152363P.

```
PR 10-SEP-1999; 99US-0153070P.
PR 13-SEP-1999; 99US-0153759P.
PR 15-SEP-1999; 99US-0154018P.
PR 16-SEP-1999; 99US-0154033P.
PR 20-SEP-1999; 99US-0154779P.
PR 22-SEP-1999; 99US-0155139P.
PR 23-SEP-1999; 99US-0155486P.
PR 24-SEP-1999; 99US-0155659P.
PR 28-SEP-1999; 99US-0156458P.
PR 29-SEP-1999; 99US-0156596P.
PR 04-OCT-1999; 99US-0157117P.
PR 05-OCT-1999; 99US-0157753P.
PR 06-OCT-1999; 99US-0157865P.
PR 07-OCT-1999; 99US-0158029P.
PR 08-OCT-1999; 99US-0158232P.
PR 12-OCT-1999; 99US-0158369P.
PR 13-OCT-1999; 99US-0159293P.
PR 13-OCT-1999; 99US-0159294P.
PR 13-OCT-1999; 99US-0159295P.
PR 14-OCT-1999; 99US-0159329P.
PR 14-OCT-1999; 99US-0159330P.
PR 14-OCT-1999; 99US-0159331P.
PR 14-OCT-1999; 99US-0159637P.
PR 14-OCT-1999; 99US-0159638P.
PR 18-OCT-1999; 99US-0159584P.
PR 21-OCT-1999; 99US-0160741P.
PR 21-OCT-1999; 99US-0160767P.
PR 21-OCT-1999; 99US-0160768P.
PR 21-OCT-1999; 99US-0160770P.
PR 21-OCT-1999; 99US-0160814P.
PR 21-OCT-1999; 99US-0160815P.
PR 22-OCT-1999; 99US-0160980P.
PR 22-OCT-1999; 99US-0160981P.
PR 22-OCT-1999; 99US-0160982P.
PR 25-OCT-1999; 99US-0161404P.
PR 25-OCT-1999; 99US-0161405P.
PR 25-OCT-1999; 99US-0161406P.
PR 26-OCT-1999; 99US-0161359P.
PR 26-OCT-1999; 99US-0161360P.
PR 26-OCT-1999; 99US-0161361P.
PR 28-OCT-1999; 99US-0161920P.
PR 28-OCT-1999; 99US-0161922P.
PR 28-OCT-1999; 99US-0161933P.
PR 29-OCT-1999; 99US-0162142P.

Query Match 100.08; Score 76; DB 3; Length 75;
Best Local Similarity 100.08; Pred. No. 4.5e-06;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 1 DEAQSKGILTLKYP 15
Db 58 DEAQSKGILTLKYP 72

RESULT 7
AAG12594
ID AAG12594 standard; protein; 79 AA.
XX
AC AAG12594;
XX
DT 17-OCT-2000 (first entry)
DE
XX Zea mays protein fragment SEQ ID NO: 11768.
XX
KW Protein identification; signal transduction pathway; metabolic pathway;
KW hybridisation assay; genetic mapping; gene expression control; promoter;
KW termination sequence; corn.
XX
OS Zea mays subsp. mays.
XX
FN EP1033405-A2.
XX
PD 06-SEP-2000.

XX 25-FEB-2000; 2000EP-00301439.
XX
PR 25-FEB-1999; 99US-0121825P.
PR 05-MAR-1999; 99US-0123180P.
PR 09-MAR-1999; 99US-0123548P.
PR 23-MAR-1999; 99US-0125788P.
PR 25-MAR-1999; 99US-0126264P.
PR 29-MAR-1999; 99US-0126785P.
PR 01-APR-1999; 99US-0127462P.
PR 06-APR-1999; 99US-0128234P.
PR 08-APR-1999; 99US-0128714P.
PR 16-APR-1999; 99US-0129845P.
PR 19-APR-1999; 99US-0130077P.
PR 21-APR-1999; 99US-0130449P.
PR 23-APR-1999; 99US-0130510P.
PR 23-APR-1999; 99US-0130891P.
PR 28-APR-1999; 99US-0131449P.
PR 30-APR-1999; 99US-0132048P.
PR 30-APR-1999; 99US-0132407P.
PR 04-MAY-1999; 99US-0132484P.
PR 05-MAY-1999; 99US-0132485P.
PR 06-MAY-1999; 99US-0132486P.
PR 06-MAY-1999; 99US-0132487P.
PR 07-MAY-1999; 99US-0132863P.
PR 11-MAY-1999; 99US-0134256P.
PR 14-MAY-1999; 99US-0134218P.
PR 14-MAY-1999; 99US-0134219P.
PR 14-MAY-1999; 99US-013421P.
PR 14-MAY-1999; 99US-0134370P.
PR 18-MAY-1999; 99US-0134768P.
PR 19-MAY-1999; 99US-0134941P.
PR 20-MAY-1999; 99US-0135124P.
PR 21-MAY-1999; 99US-0135353P.
PR 24-MAY-1999; 99US-0135629P.
PR 25-MAY-1999; 99US-0136021P.
PR 27-MAY-1999; 99US-0136392P.
PR 28-MAY-1999; 99US-0136782P.
PR 01-JUN-1999; 99US-0137222P.
PR 03-JUN-1999; 99US-0137528P.
PR 04-JUN-1999; 99US-0137502P.
PR 07-JUN-1999; 99US-0137724P.
PR 08-JUN-1999; 99US-0138094P.
PR 10-JUN-1999; 99US-0138540P.
PR 10-JUN-1999; 99US-0138847P.
PR 14-JUN-1999; 99US-0139119P.
PR 16-JUN-1999; 99US-0139452P.
PR 16-JUN-1999; 99US-0139453P.
PR 17-JUN-1999; 99US-0139492P.
PR 18-JUN-1999; 99US-0139454P.
PR 18-JUN-1999; 99US-0139455P.
PR 18-JUN-1999; 99US-0139456P.
PR 18-JUN-1999; 99US-0139457P.
PR 18-JUN-1999; 99US-0139458P.
PR 18-JUN-1999; 99US-0139459P.
PR 18-JUN-1999; 99US-0139460P.
PR 18-JUN-1999; 99US-0139461P.
PR 18-JUN-1999; 99US-0139462P.
PR 18-JUN-1999; 99US-0139463P.
PR 18-JUN-1999; 99US-0139750P.
PR 18-JUN-1999; 99US-0139763P.
PR 21-JUN-1999; 99US-0139817P.
PR 22-JUN-1999; 99US-0139899P.
PR 23-JUN-1999; 99US-0140353P.
PR 23-JUN-1999; 99US-0140354P.
PR 24-JUN-1999; 99US-0140695P.
PR 28-JUN-1999; 99US-0140823P.
PR 29-JUN-1999; 99US-0140991P.
PR 30-JUN-1999; 99US-0141287P.
PR 01-JUL-1999; 99US-0141842P.
PR 01-JUL-1999; 99US-0142154P.
PR 02-JUL-1999; 99US-0142055P.
PR 06-JUL-1999; 99US-0142390P.
```

Query Match 100.0%; Score 76; DB 3; Length 79;

PR 25-MAR-1999; 99US-0126264P.  
PR 29-MAR-1999; 99US-0126785P.  
PR 01-APR-1999; 99US-0127462P.  
PR 06-APR-1999; 99US-0128234P.  
PR 08-APR-1999; 99US-0128714P.  
PR 16-APR-1999; 99US-0129845P.  
PR 19-APR-1999; 99US-0130077P.  
PR 21-APR-1999; 99US-0130449P.  
PR 23-APR-1999; 99US-0130510P.  
PR 28-APR-1999; 99US-0130891P.  
PR 28-APR-1999; 99US-0131443P.  
PR 30-APR-1999; 99US-0132048P.  
PR 30-APR-1999; 99US-0132407P.  
PR 04-MAY-1999; 99US-0132484P.  
PR 05-MAY-1999; 99US-0132485P.  
PR 06-MAY-1999; 99US-0132486P.  
PR 06-MAY-1999; 99US-0132487P.  
PR 07-MAY-1999; 99US-0132863P.  
PR 11-MAY-1999; 99US-0134256P.  
PR 14-MAY-1999; 99US-0134218P.  
PR 14-MAY-1999; 99US-0134219P.  
PR 14-MAY-1999; 99US-0134221P.  
PR 14-MAY-1999; 99US-0134370P.  
PR 18-MAY-1999; 99US-0134768P.  
PR 19-MAY-1999; 99US-0134941P.  
PR 20-MAY-1999; 99US-0135124P.  
PR 21-MAY-1999; 99US-0135353P.  
PR 24-MAY-1999; 99US-0135623P.  
PR 25-MAY-1999; 99US-0136021P.  
PR 27-MAY-1999; 99US-0136392P.  
PR 28-MAY-1999; 99US-0136782P.  
PR 01-JUN-1999; 99US-0137222P.  
PR 03-JUN-1999; 99US-0137528P.  
PR 04-JUN-1999; 99US-0137502P.  
PR 07-JUN-1999; 99US-0137724P.  
PR 08-JUN-1999; 99US-0138094P.  
PR 10-JUN-1999; 99US-0138540P.  
PR 10-JUN-1999; 99US-0138847P.  
PR 14-JUN-1999; 99US-0139119P.  
PR 16-JUN-1999; 99US-0139452P.  
PR 16-JUN-1999; 99US-0139453P.  
PR 17-JUN-1999; 99US-0139492P.  
PR 18-JUN-1999; 99US-0139454P.  
PR 18-JUN-1999; 99US-0139455P.  
PR 18-JUN-1999; 99US-0139456P.  
PR 18-JUN-1999; 99US-0139457P.  
PR 18-JUN-1999; 99US-0139458P.  
PR 18-JUN-1999; 99US-0139459P.  
PR 18-JUN-1999; 99US-0139460P.  
PR 18-JUN-1999; 99US-0139461P.  
PR 18-JUN-1999; 99US-0139462P.  
PR 18-JUN-1999; 99US-0139463P.  
PR 18-JUN-1999; 99US-0139750P.  
PR 18-JUN-1999; 99US-0139763P.  
PR 21-JUN-1999; 99US-0139817P.  
PR 22-JUN-1999; 99US-0139899P.  
PR 23-JUN-1999; 99US-0140353P.  
PR 23-JUN-1999; 99US-0140354P.  
PR 24-JUN-1999; 99US-0140695P.  
PR 28-JUN-1999; 99US-0140823P.  
PR 29-JUN-1999; 99US-0140991P.  
PR 30-JUN-1999; 99US-0141287P.  
PR 01-JUL-1999; 99US-0141842P.  
PR 01-JUL-1999; 99US-0142154P.  
PR 02-JUL-1999; 99US-0142055P.  
PR 06-JUL-1999; 99US-0142390P.  
PR 08-JUL-1999; 99US-0142803P.  
PR 09-JUL-1999; 99US-0142920P.  
PR 12-JUL-1999; 99US-0142977P.  
PR 13-JUL-1999; 99US-0143542P.  
PR 14-JUL-1999; 99US-0143624P.  
PR 15-JUL-1999; 99US-0144005P.  
PR 16-JUL-1999; 99US-0144085P.  
PR 16-JUL-1999; 99US-0144086P.  
PR 19-JUL-1999; 99US-0144325P.  
PR 19-JUL-1999; 99US-0144331P.  
PR 19-JUL-1999; 99US-0144332P.  
PR 19-JUL-1999; 99US-0144333P.  
PR 19-JUL-1999; 99US-0144334P.  
PR 19-JUL-1999; 99US-0144335P.  
PR 20-JUL-1999; 99US-0144352P.  
PR 20-JUL-1999; 99US-0144632P.  
PR 20-JUL-1999; 99US-0144884P.  
PR 21-JUL-1999; 99US-0144814P.  
PR 21-JUL-1999; 99US-0145086P.  
PR 21-JUL-1999; 99US-0145088P.  
PR 22-JUL-1999; 99US-0145085P.  
PR 22-JUL-1999; 99US-0145087P.  
PR 22-JUL-1999; 99US-0145089P.  
PR 22-JUL-1999; 99US-0145192P.  
PR 23-JUL-1999; 99US-0145145P.  
PR 23-JUL-1999; 99US-0145218P.  
PR 23-JUL-1999; 99US-0145224P.  
PR 26-JUL-1999; 99US-0145276P.  
PR 27-JUL-1999; 99US-0145913P.  
PR 27-JUL-1999; 99US-0145918P.  
PR 27-JUL-1999; 99US-0145919P.  
PR 28-JUL-1999; 99US-0145951P.  
PR 02-AUG-1999; 99US-0146386P.  
PR 02-AUG-1999; 99US-0146388P.  
PR 02-AUG-1999; 99US-0146389P.  
PR 03-AUG-1999; 99US-0147038P.  
PR 04-AUG-1999; 99US-0147204P.  
PR 04-AUG-1999; 99US-0147302P.  
PR 05-AUG-1999; 99US-0147192P.  
PR 05-AUG-1999; 99US-0147260P.  
PR 06-AUG-1999; 99US-0147303P.  
PR 06-AUG-1999; 99US-0147416P.  
PR 09-AUG-1999; 99US-0147493P.  
PR 09-AUG-1999; 99US-0147935P.  
PR 10-AUG-1999; 99US-0148171P.  
PR 11-AUG-1999; 99US-0148319P.  
PR 12-AUG-1999; 99US-0148341P.  
PR 13-AUG-1999; 99US-0148565P.  
PR 13-AUG-1999; 99US-0148684P.  
PR 16-AUG-1999; 99US-0149368P.  
PR 17-AUG-1999; 99US-0149175P.  
PR 18-AUG-1999; 99US-0149426P.  
PR 20-AUG-1999; 99US-0149722P.  
PR 20-AUG-1999; 99US-0149723P.  
PR 20-AUG-1999; 99US-0149929P.  
PR 23-AUG-1999; 99US-0149902P.  
PR 23-AUG-1999; 99US-0149930P.  
PR 25-AUG-1999; 99US-0150566P.  
PR 26-AUG-1999; 99US-0150884P.  
PR 27-AUG-1999; 99US-0151065P.  
PR 27-AUG-1999; 99US-0151066P.  
PR 27-AUG-1999; 99US-0151080P.  
PR 30-AUG-1999; 99US-0151303P.  
PR 31-AUG-1999; 99US-0151438P.  
PR 01-SEP-1999; 99US-0151930P.  
PR 07-SEP-1999; 99US-0152363P.  
PR 10-SEP-1999; 99US-0153070P.  
PR 13-SEP-1999; 99US-0153758P.  
PR 15-SEP-1999; 99US-0154018P.  
PR 16-SEP-1999; 99US-0154039P.  
PR 20-SEP-1999; 99US-0154779P.  
PR 22-SEP-1999; 99US-0155139P.  
PR 23-SEP-1999; 99US-0155486P.  
PR 24-SEP-1999; 99US-0155659P.  
PR 28-SEP-1999; 99US-0156458P.  
PR 29-SEP-1999; 99US-0156596P.  
PR 04-OCT-1999; 99US-0157117P.  
PR 05-OCT-1999; 99US-0157753P.  
PR 06-OCT-1999; 99US-0157865P.  
PR 07-OCT-1999; 99US-0158029P.

```

PR 08-OCT-1999; 99US-0158232P.
PR 12-OCT-1999; 99US-0158369P.
PR 13-OCT-1999; 99US-0159299P.
PR 13-OCT-1999; 99US-0159299P.
PR 13-OCT-1999; 99US-0159299P.
PR 14-OCT-1999; 99US-0159329P.
PR 14-OCT-1999; 99US-0159330P.
PR 14-OCT-1999; 99US-0159331P.
PR 14-OCT-1999; 99US-0159637P.
PR 14-OCT-1999; 99US-0159638P.
PR 18-OCT-1999; 99US-0159584P.
PR 21-OCT-1999; 99US-0160741P.
PR 21-OCT-1999; 99US-0160767P.
PR 21-OCT-1999; 99US-0160768P.
PR 21-OCT-1999; 99US-0160770P.
PR 21-OCT-1999; 99US-0160814P.
PR 21-OCT-1999; 99US-0160815P.
PR 22-OCT-1999; 99US-0160980P.
PR 22-OCT-1999; 99US-0160981P.
PR 22-OCT-1999; 99US-0160982P.
PR 25-OCT-1999; 99US-0161404P.
PR 25-OCT-1999; 99US-0161405P.
PR 25-OCT-1999; 99US-0161406P.
PR 26-OCT-1999; 99US-0161359P.
PR 26-OCT-1999; 99US-0161360P.
PR 26-OCT-1999; 99US-0161361P.
PR 28-OCT-1999; 99US-0161920P.
PR 28-OCT-1999; 99US-0161922P.
PR 28-OCT-1999; 99US-0161993P.
PR 29-OCT-1999; 99US-0162142P.

Query Match 100.08; Score 76; DB 3; Length 82;
Best Local Similarity 100.08; Pred. No. 5e-06;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 1 DEAQSKRGILTLKYP 15
Db 41 DEAQSKRGILTLKYP 55

RESULT 9
AAG18916
ID AAG18916 standard; protein; 88 AA.
XX AC AAG18916;
XX AC AAG18916;
XX DT 17-OCT-2000 (first entry)
XX DE Zea mays protein fragment SEQ ID NO: 20514.
XX DE Zea mays protein fragment SEQ ID NO: 20514.
XX KW Protein identification; signal transduction pathway; metabolic pathway;
XX KW hybridisation assay; genetic mapping; gene expression control; promoter;
XX KW termination sequence; corn.
XX OS Zea mays subsp. mays.
XX OS Zea mays subsp. mays.
XX PN EP1033405-A2.
XX PD 06-SEP-2000.
XX PF 25-FEB-2000; 2000EP-00301439.
XX PF 25-FEB-1999; 99US-0121825P.
XX PF 05-MAR-1999; 99US-0123180P.
XX PF 09-MAR-1999; 99US-0123548P.
XX PF 23-MAR-1999; 99US-0125788P.
XX PF 25-MAR-1999; 99US-0126264P.
XX PF 29-MAR-1999; 99US-0126785P.
XX PF 01-APR-1999; 99US-0127462P.
XX PF 06-APR-1999; 99US-0128234P.
XX PF 08-APR-1999; 99US-0128714P.
XX PF 16-APR-1999; 99US-0129845P.
XX PF 19-APR-1999; 99US-0130077P.
PR 21-APR-1999; 99US-0130449P.
PR 23-APR-1999; 99US-0130510P.
PR 23-APR-1999; 99US-0130891P.
PR 28-APR-1999; 99US-0131449P.
PR 30-APR-1999; 99US-0132048P.
PR 30-APR-1999; 99US-0132407P.
PR 04-MAY-1999; 99US-0132484P.
PR 05-MAY-1999; 99US-0132485P.
PR 06-MAY-1999; 99US-0132486P.
PR 06-MAY-1999; 99US-0132487P.
PR 07-MAY-1999; 99US-0132863P.
PR 11-MAY-1999; 99US-0134256P.
PR 14-MAY-1999; 99US-0134218P.
PR 14-MAY-1999; 99US-0134219P.
PR 14-MAY-1999; 99US-0134221P.
PR 14-MAY-1999; 99US-0134370P.
PR 18-MAY-1999; 99US-0134768P.
PR 19-MAY-1999; 99US-0134941P.
PR 20-MAY-1999; 99US-0135124P.
PR 21-MAY-1999; 99US-0135353P.
PR 24-MAY-1999; 99US-0135629P.
PR 25-MAY-1999; 99US-0136021P.
PR 27-MAY-1999; 99US-0136392P.
PR 28-MAY-1999; 99US-0136782P.
PR 01-JUN-1999; 99US-0137222P.
PR 03-JUN-1999; 99US-0137528P.
PR 04-JUN-1999; 99US-0137502P.
PR 07-JUN-1999; 99US-0137724P.
PR 08-JUN-1999; 99US-0138094P.
PR 10-JUN-1999; 99US-0138540P.
PR 10-JUN-1999; 99US-0138847P.
PR 14-JUN-1999; 99US-0139119P.
PR 16-JUN-1999; 99US-0139452P.
PR 16-JUN-1999; 99US-0139453P.
PR 17-JUN-1999; 99US-0139492P.
PR 18-JUN-1999; 99US-0139454P.
PR 18-JUN-1999; 99US-0139455P.
PR 18-JUN-1999; 99US-0139456P.
PR 18-JUN-1999; 99US-0139457P.
PR 18-JUN-1999; 99US-0139458P.
PR 18-JUN-1999; 99US-0139459P.
PR 18-JUN-1999; 99US-0139460P.
PR 18-JUN-1999; 99US-0139461P.
PR 18-JUN-1999; 99US-0139462P.
PR 18-JUN-1999; 99US-0139463P.
PR 18-JUN-1999; 99US-0139750P.
PR 18-JUN-1999; 99US-0139763P.
PR 21-JUN-1999; 99US-0139817P.
PR 22-JUN-1999; 99US-0139899P.
PR 23-JUN-1999; 99US-0140353P.
PR 23-JUN-1999; 99US-0140354P.
PR 24-JUN-1999; 99US-0140695P.
PR 28-JUN-1999; 99US-0140823P.
PR 29-JUN-1999; 99US-0140991P.
PR 30-JUN-1999; 99US-0141287P.
PR 01-JUL-1999; 99US-0141842P.
PR 01-JUL-1999; 99US-0142154P.
PR 02-JUL-1999; 99US-0142055P.
PR 06-JUL-1999; 99US-0142390P.
PR 08-JUL-1999; 99US-0142803P.
PR 09-JUL-1999; 99US-0142920P.
PR 12-JUL-1999; 99US-0142977P.
PR 13-JUL-1999; 99US-0143542P.
PR 14-JUL-1999; 99US-0143624P.
PR 15-JUL-1999; 99US-0144005P.
PR 16-JUL-1999; 99US-0144085P.
PR 16-JUL-1999; 99US-0144086P.
PR 19-JUL-1999; 99US-0144325P.
PR 19-JUL-1999; 99US-0144331P.
PR 19-JUL-1999; 99US-0144332P.
PR 19-JUL-1999; 99US-0144333P.
PR 19-JUL-1999; 99US-0144334P.
PR 19-JUL-1999; 99US-0144335P.

```

PR 20-JUL-1999; 99US-0144352P.  
PR 20-JUL-1999; 99US-0144632P.  
PR 20-JUL-1999; 99US-0144884P.  
PR 21-JUL-1999; 99US-0144814P.  
PR 21-JUL-1999; 99US-0145086P.  
PR 21-JUL-1999; 99US-0145088P.  
PR 22-JUL-1999; 99US-0145085P.  
PR 22-JUL-1999; 99US-0145087P.  
PR 22-JUL-1999; 99US-0145089P.  
PR 22-JUL-1999; 99US-0145192P.  
PR 23-JUL-1999; 99US-0145145P.  
PR 23-JUL-1999; 99US-0145221P.  
PR 23-JUL-1999; 99US-0145224P.  
PR 26-JUL-1999; 99US-0145276P.  
PR 27-JUL-1999; 99US-0145913P.  
PR 27-JUL-1999; 99US-0145918P.  
PR 27-JUL-1999; 99US-0145919P.  
PR 28-JUL-1999; 99US-0145951P.  
PR 02-AUG-1999; 99US-0146386P.  
PR 02-AUG-1999; 99US-0146388P.  
PR 02-AUG-1999; 99US-0146389P.  
PR 03-AUG-1999; 99US-0147038P.  
PR 04-AUG-1999; 99US-0147204P.  
PR 04-AUG-1999; 99US-0147302P.  
PR 05-AUG-1999; 99US-0147192P.  
PR 05-AUG-1999; 99US-0147260P.  
PR 06-AUG-1999; 99US-0147303P.  
PR 06-AUG-1999; 99US-0147416P.  
PR 09-AUG-1999; 99US-0147493P.  
PR 09-AUG-1999; 99US-0147935P.  
PR 10-AUG-1999; 99US-0148171P.  
PR 11-AUG-1999; 99US-0148319P.  
PR 12-AUG-1999; 99US-0148341P.  
PR 13-AUG-1999; 99US-0148565P.  
PR 13-AUG-1999; 99US-0148684P.  
PR 16-AUG-1999; 99US-0149368P.  
PR 17-AUG-1999; 99US-0149175P.  
PR 18-AUG-1999; 99US-0149426P.  
PR 20-AUG-1999; 99US-0149722P.  
PR 20-AUG-1999; 99US-0149723P.  
PR 20-AUG-1999; 99US-0149923P.  
PR 23-AUG-1999; 99US-0149930P.  
PR 23-AUG-1999; 99US-0149930P.  
PR 25-AUG-1999; 99US-0150566P.  
PR 26-AUG-1999; 99US-0150884P.  
PR 27-AUG-1999; 99US-0151065P.  
PR 27-AUG-1999; 99US-0151066P.  
PR 27-AUG-1999; 99US-0151080P.  
PR 30-AUG-1999; 99US-0151303P.  
PR 31-AUG-1999; 99US-0151438P.  
PR 01-SEP-1999; 99US-0151930P.  
PR 07-SEP-1999; 99US-0152363P.  
PR 10-SEP-1999; 99US-0153070P.  
PR 13-SEP-1999; 99US-0153758P.  
PR 15-SEP-1999; 99US-0154018P.  
PR 16-SEP-1999; 99US-0154039P.  
PR 20-SEP-1999; 99US-0154779P.  
PR 22-SEP-1999; 99US-0155139P.  
PR 23-SEP-1999; 99US-0155486P.  
PR 24-SEP-1999; 99US-0155659P.  
PR 28-SEP-1999; 99US-0156458P.  
PR 29-SEP-1999; 99US-0156596P.  
PR 04-OCT-1999; 99US-0157117P.  
PR 05-OCT-1999; 99US-0157753P.  
PR 06-OCT-1999; 99US-0157865P.  
PR 07-OCT-1999; 99US-0158029P.  
PR 08-OCT-1999; 99US-0158232P.  
PR 12-OCT-1999; 99US-0158369P.  
PR 13-OCT-1999; 99US-0159293P.  
PR 13-OCT-1999; 99US-0159294P.  
PR 13-OCT-1999; 99US-0159295P.  
PR 14-OCT-1999; 99US-0159329P.  
PR 14-OCT-1999; 99US-0159330P.

PR 14-OCT-1999; 99US-0159331P.  
PR 14-OCT-1999; 99US-0159637P.  
PR 14-OCT-1999; 99US-0159638P.  
PR 18-OCT-1999; 99US-0159584P.  
PR 21-OCT-1999; 99US-0160741P.  
PR 21-OCT-1999; 99US-0160767P.  
PR 21-OCT-1999; 99US-0160768P.  
PR 21-OCT-1999; 99US-0160770P.  
PR 21-OCT-1999; 99US-0160814P.  
PR 21-OCT-1999; 99US-0160815P.  
PR 22-OCT-1999; 99US-0160980P.  
PR 22-OCT-1999; 99US-0160981P.  
PR 22-OCT-1999; 99US-0160989P.  
PR 25-OCT-1999; 99US-0161404P.  
PR 25-OCT-1999; 99US-0161405P.  
PR 25-OCT-1999; 99US-0161406P.  
PR 25-OCT-1999; 99US-0161359P.  
PR 26-OCT-1999; 99US-0161360P.  
PR 26-OCT-1999; 99US-0161361P.  
PR 28-OCT-1999; 99US-0161920P.  
PR 28-OCT-1999; 99US-0161922P.  
PR 28-OCT-1999; 99US-0161932P.  
PR 29-OCT-1999; 99US-0162142P.

Query Match 100.0%; Score 76; DB 3; Length 88;  
Best Local Similarity 100.0%; Pred. No. 5.4e-06;  
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 DEAQSKRGILTLYP 15  
| | | | | | | | | | | | | | |  
DB 13 DEAQSKRGILTLYP 27

## RESULT 10

AAG18935  
ID AAG18935 standard; protein; 91 AA.

XX AC AAG18935;

XX DT 17-OCT-2000 (first entry)

XX DE Zea mays protein fragment SEQ ID NO: 20541.

XX KW Protein identification; signal transduction pathway; metabolic pathway;  
XX KW hybridisation assay; genetic mapping; gene expression control; promoter;  
XX KW termination sequence; corn.

XX OS Zea mays subsp. mays.

XX PN EP1033405-A2.

XX PD 06-SEP-2000.

XX PF 25-FEB-2000; 2000EP-00301439.

XX XX 25-FEB-1999; 99US-0121825P.

XX XX 05-MAR-1999; 99US-0123180P.

XX XX 09-MAR-1999; 99US-0123548P.

XX XX 23-MAR-1999; 99US-0125788P.

XX XX 25-MAR-1999; 99US-0126264P.

XX XX 29-MAR-1999; 99US-0126785P.

XX XX 01-APR-1999; 99US-0127462P.

XX XX 06-APR-1999; 99US-0128234P.

XX XX 08-APR-1999; 99US-0128714P.

XX XX 16-APR-1999; 99US-0129845P.

XX XX 19-APR-1999; 99US-0130077P.

XX XX 21-APR-1999; 99US-0130449P.

XX XX 23-APR-1999; 99US-0130510P.

XX XX 23-APR-1999; 99US-0130891P.

XX XX 28-APR-1999; 99US-0131449P.

XX XX 30-APR-1999; 99US-0132048P.

XX XX 30-APR-1999; 99US-0132407P.

XX XX 04-MAY-1999; 99US-0132484P.

us-09-423-351c-9.rag

Fri Apr 8 14:14:34 2005

PR	05-MAY-1999;	99US-0132485P.
PR	06-MAY-1999;	99US-0132486P.
PR	06-MAY-1999;	99US-0132487P.
PR	07-MAY-1999;	99US-0132863P.
PR	11-MAY-1999;	99US-0134256P.
PR	14-MAY-1999;	99US-0134218P.
PR	14-MAY-1999;	99US-0134219P.
PR	14-MAY-1999;	99US-0134221P.
PR	14-MAY-1999;	99US-0134370P.
PR	18-MAY-1999;	99US-0134768P.
PR	19-MAY-1999;	99US-0134941P.
PR	20-MAY-1999;	99US-0135124P.
PR	21-MAY-1999;	99US-0135353P.
PR	24-MAY-1999;	99US-0135629P.
PR	25-MAY-1999;	99US-0136031P.
PR	27-MAY-1999;	99US-0136392P.
PR	28-MAY-1999;	99US-0136782P.
PR	01-JUN-1999;	99US-0137232P.
PR	03-JUN-1999;	99US-0137528P.
PR	04-JUN-1999;	99US-0137502P.
PR	07-JUN-1999;	99US-0137724P.
PR	08-JUN-1999;	99US-0138094P.
PR	10-JUN-1999;	99US-0138540P.
PR	10-JUN-1999;	99US-0138847P.
PR	14-JUN-1999;	99US-0139119P.
PR	16-JUN-1999;	99US-0139452P.
PR	16-JUN-1999;	99US-0139453P.
PR	17-JUN-1999;	99US-0139492P.
PR	18-JUN-1999;	99US-0139454P.
PR	18-JUN-1999;	99US-0139455P.
PR	18-JUN-1999;	99US-0139456P.
PR	18-JUN-1999;	99US-0139457P.
PR	18-JUN-1999;	99US-0139458P.
PR	18-JUN-1999;	99US-0139459P.
PR	18-JUN-1999;	99US-0139460P.
PR	18-JUN-1999;	99US-0139461P.
PR	18-JUN-1999;	99US-0139462P.
PR	18-JUN-1999;	99US-0139463P.
PR	18-JUN-1999;	99US-0139750P.
PR	18-JUN-1999;	99US-0139763P.
PR	21-JUN-1999;	99US-0139817P.
PR	22-JUN-1999;	99US-0139899P.
PR	23-JUN-1999;	99US-0140353P.
PR	23-JUN-1999;	99US-0140354P.
PR	24-JUN-1999;	99US-0140693P.
PR	28-JUN-1999;	99US-0140823P.
PR	29-JUN-1999;	99US-0140991P.
PR	30-JUN-1999;	99US-0141287P.
PR	01-JUL-1999;	99US-0141842P.
PR	01-JUL-1999;	99US-0142154P.
PR	02-JUL-1999;	99US-0142053P.
PR	06-JUL-1999;	99US-0142390P.
PR	08-JUL-1999;	99US-0142803P.
PR	09-JUL-1999;	99US-0142920P.
PR	12-JUL-1999;	99US-0142977P.
PR	13-JUL-1999;	99US-0143542P.
PR	14-JUL-1999;	99US-0143624P.
PR	15-JUL-1999;	99US-0144005P.
PR	16-JUL-1999;	99US-0144085P.
PR	16-JUL-1999;	99US-0144086P.
PR	19-JUL-1999;	99US-0144325P.
PR	19-JUL-1999;	99US-0144331P.
PR	19-JUL-1999;	99US-0144332P.
PR	19-JUL-1999;	99US-0144333P.
PR	19-JUL-1999;	99US-0144334P.
PR	19-JUL-1999;	99US-0144335P.
PR	20-JUL-1999;	99US-0144352P.
PR	20-JUL-1999;	99US-0144632P.
PR	20-JUL-1999;	99US-0144884P.
PR	21-JUL-1999;	99US-0144814P.
PR	21-JUL-1999;	99US-0145086P.
PR	21-JUL-1999;	99US-0145088P.
PR	22-JUL-1999;	99US-0145085P.
PR	22-JUL-1999;	99US-0145087P.
PR	22-JUL-1999;	99US-0145089P.
PR	22-JUL-1999;	99US-0145192P.
PR	23-JUL-1999;	99US-0145145P.
PR	23-JUL-1999;	99US-0145218P.
PR	23-JUL-1999;	99US-0145224P.
PR	26-JUL-1999;	99US-0145276P.
PR	27-JUL-1999;	99US-0145913P.
PR	27-JUL-1999;	99US-0145918P.
PR	27-JUL-1999;	99US-0145919P.
PR	27-JUL-1999;	99US-0145951P.
PR	02-AUG-1999;	99US-0146386P.
PR	02-AUG-1999;	99US-0146388P.
PR	02-AUG-1999;	99US-0146389P.
PR	03-AUG-1999;	99US-0147038P.
PR	04-AUG-1999;	99US-0147204P.
PR	04-AUG-1999;	99US-0147302P.
PR	05-AUG-1999;	99US-0147192P.
PR	05-AUG-1999;	99US-0147260P.
PR	06-AUG-1999;	99US-0147303P.
PR	08-AUG-1999;	99US-0147416P.
PR	08-AUG-1999;	99US-0147493P.
PR	09-AUG-1999;	99US-0147935P.
PR	10-AUG-1999;	99US-0148171P.
PR	11-AUG-1999;	99US-0148319P.
PR	13-AUG-1999;	99US-0148341P.
PR	13-AUG-1999;	99US-0148565P.
PR	13-AUG-1999;	99US-0148684P.
PR	16-AUG-1999;	99US-0149368P.
PR	17-AUG-1999;	99US-0149175P.
PR	18-AUG-1999;	99US-0149426P.
PR	20-AUG-1999;	99US-0149722P.
PR	20-AUG-1999;	99US-0149723P.
PR	20-AUG-1999;	99US-0149929P.
PR	23-AUG-1999;	99US-0149902P.
PR	23-AUG-1999;	99US-0149930P.
PR	25-AUG-1999;	99US-0150566P.
PR	26-AUG-1999;	99US-0150884P.
PR	27-AUG-1999;	99US-0151065P.
PR	27-AUG-1999;	99US-0151066P.
PR	30-AUG-1999;	99US-0151080P.
PR	31-AUG-1999;	99US-0151303P.
PR	01-SEP-1999;	99US-0151930P.
PR	07-SEP-1999;	99US-0152363P.
PR	10-SEP-1999;	99US-0153070P.
PR	13-SEP-1999;	99US-0153758P.
PR	15-SEP-1999;	99US-0154018P.
PR	16-SEP-1999;	99US-0154039P.
PR	22-SEP-1999;	99US-0154779P.
PR	23-SEP-1999;	99US-0155139P.
PR	24-SEP-1999;	99US-0155486P.
PR	28-SEP-1999;	99US-0155659P.
PR	29-SEP-1999;	99US-0156458P.
PR	04-OCT-1999;	99US-0156596P.
PR	05-OCT-1999;	99US-0157117P.
PR	06-OCT-1999;	99US-0157753P.
PR	07-OCT-1999;	99US-0158029P.
PR	08-OCT-1999;	99US-0158222P.
PR	12-OCT-1999;	99US-0158369P.
PR	13-OCT-1999;	99US-0159293P.
PR	13-OCT-1999;	99US-0159294P.
PR	13-OCT-1999;	99US-0159295P.
PR	14-OCT-1999;	99US-0159329P.
PR	14-OCT-1999;	99US-0159331P.
PR	14-OCT-1999;	99US-0159637P.
PR	14-OCT-1999;	99US-0159638P.
PR	18-OCT-1999;	99US-0159584P.
PR	21-OCT-1999;	99US-0160741P.
PR	21-OCT-1999;	99US-0160767P.
PR	21-OCT-1999;	99US-0160768P.



```
PR 21-OCT-1999; 99US-0160770P.
PR 21-OCT-1999; 99US-0160814P.
PR 21-OCT-1999; 99US-0160815P.
PR 22-OCT-1999; 99US-0160980P.
PR 22-OCT-1999; 99US-0160981P.
PR 22-OCT-1999; 99US-0160988P.
PR 22-OCT-1999; 99US-0160989P.
PR 25-OCT-1999; 99US-0161404P.
PR 25-OCT-1999; 99US-0161405P.
PR 25-OCT-1999; 99US-0161406P.
PR 26-OCT-1999; 99US-0161359P.
PR 26-OCT-1999; 99US-0161360P.
PR 26-OCT-1999; 99US-0161361P.
PR 28-OCT-1999; 99US-0161920P.
PR 28-OCT-1999; 99US-0161992P.
PR 28-OCT-1999; 99US-0161993P.
PR 29-OCT-1999; 99US-0162142P.

Query Match 100.0%; Score 76; DB 3; Length 91;
Best Local Similarity 100.0%; Pred. No. 5.6e-06;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 DEAQSKRGILTKYP 15
Db 58 DEAQSKRGILTKYP 72

RESULT 11
ID AAG19142
XX AAG19142 standard; protein; 95 AA.
XX AC AAG19142;
XX AD 17-OCT-2000 (first entry)
XX DT 17-OCT-2000 (first entry)
XX DE Zea mays protein fragment SEQ ID NO: 20832.
XX KW Protein identification; signal transduction pathway; metabolic pathway;
XX KW hybridisation assay; genetic mapping; gene expression control; promoter;
XX KW termination sequence; corn.
XX OS Zea mays subsp. mays.
XX PN EP1033405-A2.
XX PD 06-SEP-2000.
XX PF 25-FEB-2000; 2000EP-00301439.
XX PR 25-FEB-1999; 99US-0121825P.
XX PR 05-MAR-1999; 99US-0123180P.
XX PR 09-MAR-1999; 99US-0123548P.
XX PR 23-MAR-1999; 99US-0125788P.
XX PR 25-MAR-1999; 99US-0126264P.
XX PR 29-MAR-1999; 99US-0126785P.
XX PR 01-APR-1999; 99US-0127462P.
XX PR 06-APR-1999; 99US-0128234P.
XX PR 08-APR-1999; 99US-0128714P.
XX PR 16-APR-1999; 99US-0128845P.
XX PR 19-APR-1999; 99US-0130077P.
XX PR 21-APR-1999; 99US-0130449P.
XX PR 23-APR-1999; 99US-0130510P.
XX PR 23-APR-1999; 99US-0130891P.
XX PR 30-APR-1999; 99US-0131449P.
XX PR 30-APR-1999; 99US-0132048P.
XX PR 30-APR-1999; 99US-0132407P.
XX PR 04-MAY-1999; 99US-0132484P.
XX PR 05-MAY-1999; 99US-0132485P.
XX PR 06-MAY-1999; 99US-0132486P.
XX PR 06-MAY-1999; 99US-0132487P.
XX PR 07-MAY-1999; 99US-0132863P.
XX PR 11-MAY-1999; 99US-0134256P.
XX PR 14-MAY-1999; 99US-0134218P.
XX PR 14-MAY-1999; 99US-0134219P.
PR 14-MAY-1999; 99US-0134221P.
PR 14-MAY-1999; 99US-0134370P.
PR 18-MAY-1999; 99US-0134768P.
PR 19-MAY-1999; 99US-0134941P.
PR 20-MAY-1999; 99US-0135124P.
PR 21-MAY-1999; 99US-0135353P.
PR 24-MAY-1999; 99US-0135629P.
PR 25-MAY-1999; 99US-0136021P.
PR 27-MAY-1999; 99US-0136392P.
PR 28-MAY-1999; 99US-0136782P.
PR 01-JUN-1999; 99US-0137222P.
PR 03-JUN-1999; 99US-0137528P.
PR 04-JUN-1999; 99US-0137502P.
PR 07-JUN-1999; 99US-0137724P.
PR 08-JUN-1999; 99US-0138094P.
PR 10-JUN-1999; 99US-0138540P.
PR 10-JUN-1999; 99US-0138847P.
PR 14-JUN-1999; 99US-0139119P.
PR 16-JUN-1999; 99US-0139452P.
PR 16-JUN-1999; 99US-0139453P.
PR 17-JUN-1999; 99US-0139452P.
PR 18-JUN-1999; 99US-0139454P.
PR 18-JUN-1999; 99US-0139455P.
PR 18-JUN-1999; 99US-0139456P.
PR 18-JUN-1999; 99US-0139457P.
PR 18-JUN-1999; 99US-0139458P.
PR 18-JUN-1999; 99US-0139459P.
PR 18-JUN-1999; 99US-0139460P.
PR 18-JUN-1999; 99US-0139461P.
PR 18-JUN-1999; 99US-0139462P.
PR 18-JUN-1999; 99US-0139463P.
PR 18-JUN-1999; 99US-0139750P.
PR 18-JUN-1999; 99US-0139763P.
PR 21-JUN-1999; 99US-0139817P.
PR 22-JUN-1999; 99US-0139899P.
PR 23-JUN-1999; 99US-0140353P.
PR 24-JUN-1999; 99US-0140695P.
PR 28-JUN-1999; 99US-0140823P.
PR 29-JUN-1999; 99US-0140991P.
PR 30-JUN-1999; 99US-0141287P.
PR 01-JUL-1999; 99US-0141842P.
PR 01-JUL-1999; 99US-0142154P.
PR 02-JUL-1999; 99US-0142055P.
PR 06-JUL-1999; 99US-0142390P.
PR 08-JUL-1999; 99US-0142803P.
PR 09-JUL-1999; 99US-0142920P.
PR 12-JUL-1999; 99US-0142977P.
PR 13-JUL-1999; 99US-0143542P.
PR 14-JUL-1999; 99US-0143624P.
PR 15-JUL-1999; 99US-0144005P.
PR 16-JUL-1999; 99US-0144086P.
PR 16-JUL-1999; 99US-0144086P.
PR 19-JUL-1999; 99US-0144331P.
PR 19-JUL-1999; 99US-0144332P.
PR 19-JUL-1999; 99US-0144333P.
PR 19-JUL-1999; 99US-0144334P.
PR 19-JUL-1999; 99US-0144335P.
PR 20-JUL-1999; 99US-0144352P.
PR 20-JUL-1999; 99US-0144632P.
PR 20-JUL-1999; 99US-0144884P.
PR 21-JUL-1999; 99US-0144814P.
PR 21-JUL-1999; 99US-0145086P.
PR 21-JUL-1999; 99US-0145088P.
PR 22-JUL-1999; 99US-0145085P.
PR 22-JUL-1999; 99US-0145087P.
PR 22-JUL-1999; 99US-0145089P.
PR 22-JUL-1999; 99US-0145192P.
PR 23-JUL-1999; 99US-0145145P.
PR 23-JUL-1999; 99US-0145218P.
PR 23-JUL-1999; 99US-0145224P.
PR 26-JUL-1999; 99US-0145276P.
```

```

PR 27-JUL-1999; 99US-0145913P.
PR 27-JUL-1999; 99US-0145918P.
PR 27-JUL-1999; 99US-0145919P.
PR 28-JUL-1999; 99US-0145951P.
PR 02-AUG-1999; 99US-0146386P.
PR 02-AUG-1999; 99US-0146388P.
PR 02-AUG-1999; 99US-0146389P.
PR 03-AUG-1999; 99US-0147038P.
PR 04-AUG-1999; 99US-0147048P.
PR 04-AUG-1999; 99US-0147302P.
PR 05-AUG-1999; 99US-0147192P.
PR 05-AUG-1999; 99US-0147260P.
PR 06-AUG-1999; 99US-0147303P.
PR 06-AUG-1999; 99US-0147416P.
PR 09-AUG-1999; 99US-0147493P.
PR 09-AUG-1999; 99US-0147935P.
PR 11-AUG-1999; 99US-0148171P.
PR 11-AUG-1999; 99US-0148319P.
PR 12-AUG-1999; 99US-0148341P.
PR 13-AUG-1999; 99US-0148565P.
PR 13-AUG-1999; 99US-0148684P.
PR 16-AUG-1999; 99US-0149368P.
PR 17-AUG-1999; 99US-0149175P.
PR 18-AUG-1999; 99US-0149426P.
PR 20-AUG-1999; 99US-0149722P.
PR 20-AUG-1999; 99US-0149723P.
PR 20-AUG-1999; 99US-0149929P.
PR 23-AUG-1999; 99US-0149902P.
PR 23-AUG-1999; 99US-0149930P.
PR 25-AUG-1999; 99US-0150566P.
PR 26-AUG-1999; 99US-0150884P.
PR 27-AUG-1999; 99US-0151065P.
PR 27-AUG-1999; 99US-0151066P.
PR 27-AUG-1999; 99US-0151080P.
PR 30-AUG-1999; 99US-0151303P.
PR 31-AUG-1999; 99US-0151438P.
PR 01-SEP-1999; 99US-0151930P.
PR 07-SEP-1999; 99US-0152363P.
PR 10-SEP-1999; 99US-0153070P.
PR 13-SEP-1999; 99US-0153758P.
PR 15-SEP-1999; 99US-0154018P.
PR 16-SEP-1999; 99US-0154039P.
PR 20-SEP-1999; 99US-0154779P.
PR 22-SEP-1999; 99US-0155139P.
PR 23-SEP-1999; 99US-0155486P.
PR 24-SEP-1999; 99US-0155659P.
PR 28-SEP-1999; 99US-0156458P.
PR 29-SEP-1999; 99US-0156596P.
PR 04-OCT-1999; 99US-0157117P.
PR 05-OCT-1999; 99US-0157753P.
PR 06-OCT-1999; 99US-0157865P.
PR 07-OCT-1999; 99US-0158029P.
PR 08-OCT-1999; 99US-0158232P.
PR 12-OCT-1999; 99US-0158369P.
PR 13-OCT-1999; 99US-0159294P.
PR 13-OCT-1999; 99US-0159298P.
PR 14-OCT-1999; 99US-0159329P.
PR 14-OCT-1999; 99US-0159330P.
PR 14-OCT-1999; 99US-0159331P.
PR 14-OCT-1999; 99US-0159637P.
PR 14-OCT-1999; 99US-0159638P.
PR 18-OCT-1999; 99US-0159584P.
PR 21-OCT-1999; 99US-0160741P.
PR 21-OCT-1999; 99US-0160767P.
PR 21-OCT-1999; 99US-0160768P.
PR 21-OCT-1999; 99US-0160770P.
PR 21-OCT-1999; 99US-0160814P.
PR 21-OCT-1999; 99US-0160815P.
PR 22-OCT-1999; 99US-0160980P.
PR 22-OCT-1999; 99US-0160981P.
PR 22-OCT-1999; 99US-0160989P.
PR 25-OCT-1999; 99US-0161404P.

PR 25-OCT-1999; 99US-0161405P.
PR 25-OCT-1999; 99US-0161406P.
PR 26-OCT-1999; 99US-0161359P.
PR 26-OCT-1999; 99US-0161360P.
PR 26-OCT-1999; 99US-0161361P.
PR 28-OCT-1999; 99US-0161920P.
PR 28-OCT-1999; 99US-0161922P.
PR 28-OCT-1999; 99US-0161993P.
PR 28-OCT-1999; 99US-0162142P.

Query Match 100.0%; Score 76; DB 3; Length 95;
Best Local Similarity 100.0%; Pred. No. 5.9e-06;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 DEAQSKRGILTLKYP 15
Db 41 DEAQSKRGILTLKYP 55

RESULT 12
AAG12593
ID AAG12593 standard; protein; 96 AA.
XX
AC AAG12593;
XX
DT 17-OCT-2000 (first entry)
XX
DE Zea mays protein fragment SEQ ID NO: 11767.
XX
KW Protein identification; signal transduction pathway; metabolic pathway;
KW hybridisation assay; genetic mapping; gene expression control; promoter;
KW termination sequence; corn.
XX
OS Zea mays subsp. mays.
XX
PN EP1033405-A2.
XX
PD 06-SEP-2000.
XX
PF 25-FEB-2000; 2000EP-00301439.
XX
PR 25-FEB-1999; 99US-0121825P.
PR 05-MAR-1999; 99US-0123180P.
PR 09-MAR-1999; 99US-0123548P.
PR 23-MAR-1999; 99US-0125788P.
PR 25-MAR-1999; 99US-0126264P.
PR 29-MAR-1999; 99US-0126785P.
PR 01-APR-1999; 99US-0127462P.
PR 06-APR-1999; 99US-0128234P.
PR 16-APR-1999; 99US-0128714P.
PR 08-APR-1999; 99US-0129845P.
PR 19-APR-1999; 99US-0130077P.
PR 21-APR-1999; 99US-0130449P.
PR 23-APR-1999; 99US-0130510P.
PR 23-APR-1999; 99US-0130891P.
PR 28-APR-1999; 99US-0131449P.
PR 30-APR-1999; 99US-0132048P.
PR 30-APR-1999; 99US-0132407P.
PR 04-MAY-1999; 99US-0132484P.
PR 05-MAY-1999; 99US-0132485P.
PR 06-MAY-1999; 99US-0132486P.
PR 06-MAY-1999; 99US-0132487P.
PR 07-MAY-1999; 99US-0132863P.
PR 11-MAY-1999; 99US-0134256P.
PR 14-MAY-1999; 99US-0134218P.
PR 14-MAY-1999; 99US-0134219P.
PR 14-MAY-1999; 99US-0134221P.
PR 14-MAY-1999; 99US-0134370P.
PR 18-MAY-1999; 99US-0134768P.
PR 19-MAY-1999; 99US-0134941P.
PR 20-MAY-1999; 99US-0135124P.
PR 21-MAY-1999; 99US-0135353P.
PR 24-MAY-1999; 99US-0135629P.

```

PR 25-MAY-1999; 99US-0136021P.  
PR 27-MAY-1999; 99US-0136392P.  
PR 28-MAY-1999; 99US-0136782P.  
PR 01-JUN-1999; 99US-0137222P.  
PR 03-JUN-1999; 99US-0137528P.  
PR 04-JUN-1999; 99US-0137503P.  
PR 07-JUN-1999; 99US-0137724P.  
PR 08-JUN-1999; 99US-0138094P.  
PR 10-JUN-1999; 99US-0138540P.  
PR 10-JUN-1999; 99US-0138847P.  
PR 14-JUN-1999; 99US-0139119P.  
PR 16-JUN-1999; 99US-0139452P.  
PR 16-JUN-1999; 99US-0139453P.  
PR 17-JUN-1999; 99US-0139492P.  
PR 18-JUN-1999; 99US-0139454P.  
PR 18-JUN-1999; 99US-0139455P.  
PR 18-JUN-1999; 99US-0139456P.  
PR 18-JUN-1999; 99US-0139457P.  
PR 18-JUN-1999; 99US-0139458P.  
PR 18-JUN-1999; 99US-0139459P.  
PR 18-JUN-1999; 99US-0139460P.  
PR 18-JUN-1999; 99US-0139461P.  
PR 18-JUN-1999; 99US-0139462P.  
PR 18-JUN-1999; 99US-0139463P.  
PR 18-JUN-1999; 99US-0139750P.  
PR 18-JUN-1999; 99US-0139763P.  
PR 21-JUN-1999; 99US-0139817P.  
PR 22-JUN-1999; 99US-0139899P.  
PR 23-JUN-1999; 99US-0140353P.  
PR 23-JUN-1999; 99US-0140354P.  
PR 24-JUN-1999; 99US-0140695P.  
PR 28-JUN-1999; 99US-0140823P.  
PR 29-JUN-1999; 99US-0140991P.  
PR 30-JUN-1999; 99US-0141287P.  
PR 01-JUL-1999; 99US-0141842P.  
PR 01-JUL-1999; 99US-0142154P.  
PR 02-JUL-1999; 99US-0142055P.  
PR 06-JUL-1999; 99US-0142390P.  
PR 08-JUL-1999; 99US-0142803P.  
PR 09-JUL-1999; 99US-0142920P.  
PR 12-JUL-1999; 99US-0142977P.  
PR 13-JUL-1999; 99US-0143542P.  
PR 14-JUL-1999; 99US-0143624P.  
PR 15-JUL-1999; 99US-0144005P.  
PR 16-JUL-1999; 99US-0144085P.  
PR 16-JUL-1999; 99US-0144086P.  
PR 19-JUL-1999; 99US-0144325P.  
PR 19-JUL-1999; 99US-0144331P.  
PR 19-JUL-1999; 99US-0144332P.  
PR 19-JUL-1999; 99US-0144333P.  
PR 19-JUL-1999; 99US-0144334P.  
PR 19-JUL-1999; 99US-0144335P.  
PR 20-JUL-1999; 99US-0144352P.  
PR 20-JUL-1999; 99US-0144632P.  
PR 20-JUL-1999; 99US-0144884P.  
PR 21-JUL-1999; 99US-0144814P.  
PR 21-JUL-1999; 99US-0145086P.  
PR 21-JUL-1999; 99US-0145088P.  
PR 22-JUL-1999; 99US-0145085P.  
PR 22-JUL-1999; 99US-0145087P.  
PR 22-JUL-1999; 99US-0145089P.  
PR 22-JUL-1999; 99US-0145192P.  
PR 23-JUL-1999; 99US-014514P.  
PR 23-JUL-1999; 99US-0145218P.  
PR 23-JUL-1999; 99US-0145224P.  
PR 26-JUL-1999; 99US-0145276P.  
PR 27-JUL-1999; 99US-0145913P.  
PR 27-JUL-1999; 99US-0145918P.  
PR 27-JUL-1999; 99US-0145919P.  
PR 28-JUL-1999; 99US-0145951P.  
PR 02-AUG-1999; 99US-0146386P.  
PR 02-AUG-1999; 99US-0146388P.  
PR 02-AUG-1999; 99US-0146389P.  
PR 02-AUG-1999; 99US-0146389P.  
PR 03-AUG-1999; 99US-0147038P.  
PR 04-AUG-1999; 99US-0147204P.  
PR 04-AUG-1999; 99US-0147302P.  
PR 05-AUG-1999; 99US-0147192P.  
PR 05-AUG-1999; 99US-0147260P.  
PR 06-AUG-1999; 99US-0147303P.  
PR 06-AUG-1999; 99US-0147416P.  
PR 09-AUG-1999; 99US-0147493P.  
PR 09-AUG-1999; 99US-0147935P.  
PR 10-AUG-1999; 99US-0148171P.  
PR 11-AUG-1999; 99US-0148319P.  
PR 12-AUG-1999; 99US-0148341P.  
PR 13-AUG-1999; 99US-0148565P.  
PR 13-AUG-1999; 99US-0148684P.  
PR 16-AUG-1999; 99US-0149368P.  
PR 17-AUG-1999; 99US-0149175P.  
PR 18-AUG-1999; 99US-0149436P.  
PR 20-AUG-1999; 99US-0149722P.  
PR 20-AUG-1999; 99US-0149723P.  
PR 20-AUG-1999; 99US-0149929P.  
PR 23-AUG-1999; 99US-0149902P.  
PR 23-AUG-1999; 99US-0149930P.  
PR 25-AUG-1999; 99US-0150566P.  
PR 26-AUG-1999; 99US-0150884P.  
PR 27-AUG-1999; 99US-0151065P.  
PR 27-AUG-1999; 99US-0151066P.  
PR 27-AUG-1999; 99US-0151080P.  
PR 30-AUG-1999; 99US-0151303P.  
PR 31-AUG-1999; 99US-0151438P.  
PR 01-SEP-1999; 99US-0151930P.  
PR 07-SEP-1999; 99US-0152363P.  
PR 10-SEP-1999; 99US-0153070P.  
PR 13-SEP-1999; 99US-0153758P.  
PR 15-SEP-1999; 99US-0154018P.  
PR 16-SEP-1999; 99US-0154039P.  
PR 20-SEP-1999; 99US-0154779P.  
PR 22-SEP-1999; 99US-0155139P.  
PR 23-SEP-1999; 99US-0155486P.  
PR 24-SEP-1999; 99US-0155659P.  
PR 28-SEP-1999; 99US-0156458P.  
PR 29-SEP-1999; 99US-0156596P.  
PR 04-OCT-1999; 99US-0157117P.  
PR 05-OCT-1999; 99US-015753P.  
PR 06-OCT-1999; 99US-0157865P.  
PR 07-OCT-1999; 99US-0158029P.  
PR 08-OCT-1999; 99US-0158232P.  
PR 12-OCT-1999; 99US-0158369P.  
PR 13-OCT-1999; 99US-0159293P.  
PR 13-OCT-1999; 99US-0159294P.  
PR 13-OCT-1999; 99US-0159295P.  
PR 14-OCT-1999; 99US-0159329P.  
PR 14-OCT-1999; 99US-0159330P.  
PR 14-OCT-1999; 99US-0159331P.  
PR 14-OCT-1999; 99US-0159637P.  
PR 18-OCT-1999; 99US-0159638P.  
PR 18-OCT-1999; 99US-0159584P.  
PR 21-OCT-1999; 99US-0160741P.  
PR 21-OCT-1999; 99US-0160767P.  
PR 21-OCT-1999; 99US-0160768P.  
PR 21-OCT-1999; 99US-0160770P.  
PR 21-OCT-1999; 99US-0160814P.  
PR 21-OCT-1999; 99US-0160815P.  
PR 22-OCT-1999; 99US-0160980P.  
PR 22-OCT-1999; 99US-0160981P.  
PR 22-OCT-1999; 99US-0160989P.  
PR 25-OCT-1999; 99US-0161404P.  
PR 25-OCT-1999; 99US-0161405P.  
PR 25-OCT-1999; 99US-0161406P.  
PR 26-OCT-1999; 99US-0161359P.  
PR 26-OCT-1999; 99US-0161360P.  
PR 26-OCT-1999; 99US-0161361P.  
PR 28-OCT-1999; 99US-0161920P.  
PR 28-OCT-1999; 99US-0161922P.

```
PR 28-OCT-1999; 99US-0161993P.
PR 29-OCT-1999; 99US-0162142P.

Query Match      100.08; Score 76; DB 3; Length 96;
Best Local Similarity 100.08; Pred. No. 5.9e-06;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 1 DEAQSKRGILTLYKP 15
    |||||
Db 58 DEAQSKRGILTLYKP 72

RESULT 13
AAG18903
ID AAG18903 standard; protein; 99 AA.
XX AC AAG18903;
XX DT 17-OCT-2000 (first entry)
XX DE Zea mays protein fragment SEQ ID NO: 20497.
XX KW Protein identification; signal transduction pathway; metabolic pathway;
KW hybridisation assay; genetic mapping; gene expression control; promoter;
XX termination sequence; corn.
XX OS Zea mays subsp. mays.
XX PN EP1033405-A2.
XX PD 06-SEP-2000.
XX PF 25-FEB-2000; 2000EP-00301439.
XX PR 25-FEB-1999; 99US-0121825P.
PR 05-MAR-1999; 99US-0123180P.
PR 09-MAR-1999; 99US-0123548P.
PR 23-MAR-1999; 99US-0125788P.
PR 25-MAR-1999; 99US-0126264P.
PR 29-MAR-1999; 99US-0126785P.
PR 01-APR-1999; 99US-0127462P.
PR 06-APR-1999; 99US-0128234P.
PR 08-APR-1999; 99US-0128714P.
PR 16-APR-1999; 99US-0129845P.
PR 19-APR-1999; 99US-0130077P.
PR 21-APR-1999; 99US-0130449P.
PR 23-APR-1999; 99US-0130510P.
PR 28-APR-1999; 99US-0130891P.
PR 30-APR-1999; 99US-0131449P.
PR 30-APR-1999; 99US-0132048P.
PR 30-APR-1999; 99US-0132407P.
PR 04-MAY-1999; 99US-0132484P.
PR 05-MAY-1999; 99US-0132485P.
PR 06-MAY-1999; 99US-0132486P.
PR 06-MAY-1999; 99US-0132487P.
PR 07-MAY-1999; 99US-0132863P.
PR 11-MAY-1999; 99US-0134256P.
PR 14-MAY-1999; 99US-0134218P.
PR 14-MAY-1999; 99US-0134219P.
PR 14-MAY-1999; 99US-0134221P.
PR 14-MAY-1999; 99US-0134370P.
PR 18-MAY-1999; 99US-0134768P.
PR 19-MAY-1999; 99US-0134941P.
PR 20-MAY-1999; 99US-0135124P.
PR 21-MAY-1999; 99US-0135353P.
PR 24-MAY-1999; 99US-0135629P.
PR 25-MAY-1999; 99US-0136021P.
PR 27-MAY-1999; 99US-0136392P.
PR 28-MAY-1999; 99US-0136782P.
PR 01-JUN-1999; 99US-0137222P.
PR 03-JUN-1999; 99US-0137528P.
PR 04-JUN-1999; 99US-0137502P.
PR 07-JUN-1999; 99US-0137724P.

08-JUN-1999; 99US-0138094P.
10-JUN-1999; 99US-0138540P.
10-JUN-1999; 99US-0138847P.
14-JUN-1999; 99US-0139119P.
16-JUN-1999; 99US-0139452P.
16-JUN-1999; 99US-0139453P.
17-JUN-1999; 99US-0139492P.
18-JUN-1999; 99US-0139454P.
18-JUN-1999; 99US-0139455P.
18-JUN-1999; 99US-0139456P.
18-JUN-1999; 99US-0139457P.
18-JUN-1999; 99US-0139458P.
18-JUN-1999; 99US-0139459P.
18-JUN-1999; 99US-0139461P.
18-JUN-1999; 99US-0139462P.
18-JUN-1999; 99US-0139463P.
18-JUN-1999; 99US-0139750P.
18-JUN-1999; 99US-0139763P.
21-JUN-1999; 99US-0139817P.
22-JUN-1999; 99US-0139859P.
23-JUN-1999; 99US-0140353P.
23-JUN-1999; 99US-0140354P.
24-JUN-1999; 99US-0140695P.
28-JUN-1999; 99US-0140823P.
29-JUN-1999; 99US-0140991P.
30-JUN-1999; 99US-0141287P.
01-JUL-1999; 99US-0141842P.
01-JUL-1999; 99US-0142154P.
02-JUL-1999; 99US-0142055P.
06-JUL-1999; 99US-0142390P.
08-JUL-1999; 99US-0142803P.
09-JUL-1999; 99US-0142920P.
12-JUL-1999; 99US-0142977P.
13-JUL-1999; 99US-0143542P.
14-JUL-1999; 99US-0143624P.
15-JUL-1999; 99US-0144005P.
16-JUL-1999; 99US-0144085P.
16-JUL-1999; 99US-0144086P.
19-JUL-1999; 99US-0144325P.
19-JUL-1999; 99US-0144331P.
19-JUL-1999; 99US-0144332P.
19-JUL-1999; 99US-0144333P.
19-JUL-1999; 99US-0144334P.
19-JUL-1999; 99US-0144335P.
20-JUL-1999; 99US-0144352P.
20-JUL-1999; 99US-0144632P.
20-JUL-1999; 99US-0144884P.
21-JUL-1999; 99US-0144814P.
21-JUL-1999; 99US-0145086P.
21-JUL-1999; 99US-0145088P.
22-JUL-1999; 99US-0145085P.
22-JUL-1999; 99US-0145087P.
22-JUL-1999; 99US-0145089P.
22-JUL-1999; 99US-0145192P.
23-JUL-1999; 99US-0145145P.
23-JUL-1999; 99US-0145218P.
23-JUL-1999; 99US-0145224P.
26-JUL-1999; 99US-0145276P.
27-JUL-1999; 99US-0145913P.
27-JUL-1999; 99US-0145918P.
27-JUL-1999; 99US-0145919P.
28-JUL-1999; 99US-0145951P.
02-AUG-1999; 99US-0146386P.
02-AUG-1999; 99US-0146388P.
02-AUG-1999; 99US-0146389P.
03-AUG-1999; 99US-0147038P.
04-AUG-1999; 99US-0147204P.
04-AUG-1999; 99US-0147302P.
05-AUG-1999; 99US-014732P.
05-AUG-1999; 99US-0147260P.
06-AUG-1999; 99US-0147303P.
06-AUG-1999; 99US-0147416P.
```

PR	09-AUG-1999;	99US-0147493P.
PR	09-AUG-1999;	99US-0147935P.
PR	10-AUG-1999;	99US-0148171P.
PR	11-AUG-1999;	99US-0148319P.
PR	12-AUG-1999;	99US-0148341P.
PR	13-AUG-1999;	99US-0148565P.
PR	13-AUG-1999;	99US-0148565P.
PR	16-AUG-1999;	99US-0148684P.
PR	17-AUG-1999;	99US-0149175P.
PR	18-AUG-1999;	99US-0149426P.
PR	20-AUG-1999;	99US-0149722P.
PR	20-AUG-1999;	99US-0149723P.
PR	20-AUG-1999;	99US-0149929P.
PR	23-AUG-1999;	99US-0149902P.
PR	23-AUG-1999;	99US-0149930P.
PR	25-AUG-1999;	99US-0150566P.
PR	26-AUG-1999;	99US-0150884P.
PR	27-AUG-1999;	99US-015065P.
PR	27-AUG-1999;	99US-0151066P.
PR	27-AUG-1999;	99US-0151080P.
PR	30-AUG-1999;	99US-0151303P.
PR	31-AUG-1999;	99US-0151438P.
PR	01-SEP-1999;	99US-0151930P.
PR	07-SEP-1999;	99US-0152363P.
PR	10-SEP-1999;	99US-0153070P.
PR	13-SEP-1999;	99US-0153758P.
PR	15-SEP-1999;	99US-0154018P.
PR	16-SEP-1999;	99US-0154039P.
PR	20-SEP-1999;	99US-0154779P.
PR	22-SEP-1999;	99US-0155139P.
PR	23-SEP-1999;	99US-0155486P.
PR	24-SEP-1999;	99US-0155659P.
PR	28-SEP-1999;	99US-0156458P.
PR	29-SEP-1999;	99US-0156598P.
PR	04-OCT-1999;	99US-0157117P.
PR	05-OCT-1999;	99US-0157753P.
PR	06-OCT-1999;	99US-0157865P.
PR	07-OCT-1999;	99US-0158023P.
PR	08-OCT-1999;	99US-0158232P.
PR	12-OCT-1999;	99US-0158369P.
PR	13-OCT-1999;	99US-0159293P.
PR	13-OCT-1999;	99US-0159294P.
PR	14-OCT-1999;	99US-0159295P.
PR	14-OCT-1999;	99US-0159330P.
PR	14-OCT-1999;	99US-0159331P.
PR	14-OCT-1999;	99US-0159637P.
PR	14-OCT-1999;	99US-0159638P.
PR	21-OCT-1999;	99US-0159584P.
PR	21-OCT-1999;	99US-0160741P.
PR	21-OCT-1999;	99US-0160767P.
PR	21-OCT-1999;	99US-0160768P.
PR	21-OCT-1999;	99US-0160770P.
PR	21-OCT-1999;	99US-0160814P.
PR	21-OCT-1999;	99US-0160815P.
PR	22-OCT-1999;	99US-0160980P.
PR	22-OCT-1999;	99US-0160981P.
PR	22-OCT-1999;	99US-0160989P.
PR	25-OCT-1999;	99US-0161404P.
PR	25-OCT-1999;	99US-0161405P.
PR	25-OCT-1999;	99US-0161406P.
PR	26-OCT-1999;	99US-0161359P.
PR	26-OCT-1999;	99US-0161360P.
PR	26-OCT-1999;	99US-0161361P.
PR	28-OCT-1999;	99US-0161920P.
PR	28-OCT-1999;	99US-0161992P.
PR	28-OCT-1999;	99US-0161993P.
PR	29-OCT-1999;	99US-0162142P.

Query Match 100.0%; Score 76; DB 3; Length 99;  
Beat Local Similarity 100.0%; Pred. No. 6.1e-06;  
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QV	1	DEAQRGILTKYP	15
Db	58	DEAQRGILTKYP	72
RESULT 14			
AAG27056			
ID	AAG27056 standard; protein; 101 AA.		
XX			
AC	AAG27056;		
XX			
DT	17-OCT-2000 (first entry)		
XX			
DE	Zea mays protein fragment SEQ ID NO:		
XX			
KW	Protein identification; signal transduction pathway; metabolic pathway;		
KW	hybridisation assay; genetic mapping		
KW	termination sequence; corn.		
XX			
OS	Zea mays subsp. mays.		
XX			
PN	EP1033405-A2.		
XX			
PD	06-SEP-2000.		
XX			
PF	25-FEB-2000; 2000EP-00301439.		
XX			
PR	25-FEB-1999;	99US-0121825P.	
PR	05-MAR-1999;	99US-0123180P.	
PR	09-MAR-1999;	99US-0123548P.	
PR	23-MAR-1999;	99US-0125788P.	
PR	25-MAR-1999;	99US-0126264P.	
PR	29-MAR-1999;	99US-0126785P.	
PR	01-APR-1999;	99US-0127462P.	
PR	06-APR-1999;	99US-0128234P.	
PR	08-APR-1999;	99US-0128714P.	
PR	16-APR-1999;	99US-0129845P.	
PR	19-APR-1999;	99US-0130077P.	
PR	21-APR-1999;	99US-0130449P.	
PR	23-APR-1999;	99US-0130510P.	
PR	23-APR-1999;	99US-0130891P.	
PR	28-APR-1999;	99US-0131449P.	
PR	30-APR-1999;	99US-0132048P.	
PR	30-APR-1999;	99US-0132407P.	
PR	04-MAY-1999;	99US-0132484P.	
PR	05-MAY-1999;	99US-0132485P.	
PR	06-MAY-1999;	99US-0132486P.	
PR	06-MAY-1999;	99US-0132487P.	
PR	07-MAY-1999;	99US-0132863P.	
PR	11-MAY-1999;	99US-0134256P.	
PR	14-MAY-1999;	99US-0134218P.	
PR	14-MAY-1999;	99US-0134219P.	
PR	14-MAY-1999;	99US-0134221P.	
PR	14-MAY-1999;	99US-0134370P.	
PR	18-MAY-1999;	99US-0134768P.	
PR	19-MAY-1999;	99US-0134941P.	
PR	20-MAY-1999;	99US-0135124P.	
PR	21-MAY-1999;	99US-0135353P.	
PR	24-MAY-1999;	99US-0135629P.	
PR	25-MAY-1999;	99US-0136021P.	
PR	27-MAY-1999;	99US-0136392P.	
PR	28-MAY-1999;	99US-0136782P.	
PR	01-JUN-1999;	99US-0137222P.	
PR	03-JUN-1999;	99US-0137528P.	
PR	04-JUN-1999;	99US-0137502P.	
PR	07-JUN-1999;	99US-0137724P.	
PR	08-JUN-1999;	99US-0138094P.	
PR	10-JUN-1999;	99US-0138540P.	
PR	10-JUN-1999;	99US-0138847P.	
PR	14-JUN-1999;	99US-0139119P.	
PR	16-JUN-1999;	99US-0139452P.	
PR	16-JUN-1999;	99US-0139453P.	
PR	17-JUN-1999;	99US-0139492P.	
PR			

PR 18-JUN-1999; 99US-0139454P.  
PR 18-JUN-1999; 99US-0139455P.  
PR 18-JUN-1999; 99US-0139456P.  
PR 18-JUN-1999; 99US-0139457P.  
PR 18-JUN-1999; 99US-0139458P.  
PR 18-JUN-1999; 99US-0139459P.  
PR 18-JUN-1999; 99US-0139460P.  
PR 18-JUN-1999; 99US-0139461P.  
PR 18-JUN-1999; 99US-0139462P.  
PR 18-JUN-1999; 99US-0139463P.  
PR 18-JUN-1999; 99US-0139763P.  
PR 21-JUN-1999; 99US-0139817P.  
PR 22-JUN-1999; 99US-0139899P.  
PR 23-JUN-1999; 99US-0140353P.  
PR 23-JUN-1999; 99US-0140354P.  
PR 24-JUN-1999; 99US-0140695P.  
PR 28-JUN-1999; 99US-0140823P.  
PR 29-JUN-1999; 99US-014091P.  
PR 30-JUN-1999; 99US-0141287P.  
PR 01-JUL-1999; 99US-0141842P.  
PR 01-JUL-1999; 99US-0142154P.  
PR 02-JUL-1999; 99US-0142055P.  
PR 06-JUL-1999; 99US-0142390P.  
PR 08-JUL-1999; 99US-0142803P.  
PR 09-JUL-1999; 99US-0142920P.  
PR 12-JUL-1999; 99US-0142977P.  
PR 13-JUL-1999; 99US-0143542P.  
PR 14-JUL-1999; 99US-0143624P.  
PR 15-JUL-1999; 99US-0144005P.  
PR 16-JUL-1999; 99US-0144085P.  
PR 16-JUL-1999; 99US-0144086P.  
PR 19-JUL-1999; 99US-0144332P.  
PR 19-JUL-1999; 99US-0144331P.  
PR 19-JUL-1999; 99US-0144333P.  
PR 19-JUL-1999; 99US-0144334P.  
PR 19-JUL-1999; 99US-0144335P.  
PR 20-JUL-1999; 99US-0144332P.  
PR 20-JUL-1999; 99US-0144632P.  
PR 20-JUL-1999; 99US-0144884P.  
PR 21-JUL-1999; 99US-0144814P.  
PR 21-JUL-1999; 99US-0145086P.  
PR 21-JUL-1999; 99US-0145088P.  
PR 22-JUL-1999; 99US-0145085P.  
PR 22-JUL-1999; 99US-0145087P.  
PR 22-JUL-1999; 99US-0145089P.  
PR 22-JUL-1999; 99US-0145192P.  
PR 23-JUL-1999; 99US-0145145P.  
PR 23-JUL-1999; 99US-0145218P.  
PR 23-JUL-1999; 99US-0145224P.  
PR 26-JUL-1999; 99US-0145276P.  
PR 27-JUL-1999; 99US-0145913P.  
PR 27-JUL-1999; 99US-0145918P.  
PR 27-JUL-1999; 99US-0145919P.  
PR 28-JUL-1999; 99US-0145951P.  
PR 02-AUG-1999; 99US-0146386P.  
PR 02-AUG-1999; 99US-0146388P.  
PR 02-AUG-1999; 99US-0146389P.  
PR 03-AUG-1999; 99US-0147038P.  
PR 04-AUG-1999; 99US-0147204P.  
PR 04-AUG-1999; 99US-0147302P.  
PR 05-AUG-1999; 99US-0147192P.  
PR 05-AUG-1999; 99US-0147260P.  
PR 06-AUG-1999; 99US-0147303P.  
PR 06-AUG-1999; 99US-0147416P.  
PR 09-AUG-1999; 99US-0147493P.  
PR 09-AUG-1999; 99US-0147935P.  
PR 10-AUG-1999; 99US-0148171P.  
PR 11-AUG-1999; 99US-0148319P.  
PR 12-AUG-1999; 99US-0148341P.  
PR 13-AUG-1999; 99US-0148565P.  
PR 13-AUG-1999; 99US-0148684P.

PR 16-AUG-1999; 99US-0149368P.  
PR 17-AUG-1999; 99US-0149175P.  
PR 18-AUG-1999; 99US-0149426P.  
PR 20-AUG-1999; 99US-0149722P.  
PR 20-AUG-1999; 99US-0149723P.  
PR 20-AUG-1999; 99US-0149929P.  
PR 23-AUG-1999; 99US-0149902P.  
PR 23-AUG-1999; 99US-0149930P.  
PR 25-AUG-1999; 99US-0150566P.  
PR 25-AUG-1999; 99US-0150884P.  
PR 26-AUG-1999; 99US-0151065P.  
PR 27-AUG-1999; 99US-0151066P.  
PR 27-AUG-1999; 99US-0151080P.  
PR 30-AUG-1999; 99US-0151303P.  
PR 31-AUG-1999; 99US-0151438P.  
PR 01-SEP-1999; 99US-0151930P.  
PR 07-SEP-1999; 99US-0152363P.  
PR 10-SEP-1999; 99US-0153070P.  
PR 13-SEP-1999; 99US-0153758P.  
PR 13-SEP-1999; 99US-0154018P.  
PR 16-SEP-1999; 99US-0154039P.  
PR 20-SEP-1999; 99US-0154779P.  
PR 22-SEP-1999; 99US-0155139P.  
PR 23-SEP-1999; 99US-0155486P.  
PR 24-SEP-1999; 99US-0155659P.  
PR 28-SEP-1999; 99US-0156458P.  
PR 28-SEP-1999; 99US-0156596P.  
PR 04-OCT-1999; 99US-0157117P.  
PR 06-OCT-1999; 99US-0157753P.  
PR 07-OCT-1999; 99US-0157865P.  
PR 08-OCT-1999; 99US-0158232P.  
PR 12-OCT-1999; 99US-0158369P.  
PR 13-OCT-1999; 99US-0159293P.  
PR 13-OCT-1999; 99US-0159294P.  
PR 13-OCT-1999; 99US-0159295P.  
PR 14-OCT-1999; 99US-0159329P.  
PR 14-OCT-1999; 99US-0159330P.  
PR 14-OCT-1999; 99US-0159331P.  
PR 14-OCT-1999; 99US-0159637P.  
PR 18-OCT-1999; 99US-0159638P.  
PR 18-OCT-1999; 99US-0160741P.  
PR 21-OCT-1999; 99US-0160767P.  
PR 21-OCT-1999; 99US-0160768P.  
PR 21-OCT-1999; 99US-0160770P.  
PR 21-OCT-1999; 99US-0160814P.  
PR 21-OCT-1999; 99US-0160815P.  
PR 22-OCT-1999; 99US-0160980P.  
PR 22-OCT-1999; 99US-0160981P.  
PR 22-OCT-1999; 99US-0160989P.  
PR 25-OCT-1999; 99US-0161404P.  
PR 25-OCT-1999; 99US-0161405P.  
PR 25-OCT-1999; 99US-0161406P.  
PR 26-OCT-1999; 99US-0161359P.  
PR 26-OCT-1999; 99US-0161360P.  
PR 26-OCT-1999; 99US-0161361P.  
PR 28-OCT-1999; 99US-0161920P.  
PR 28-OCT-1999; 99US-0161992P.  
PR 28-OCT-1999; 99US-0161993P.  
PR 29-OCT-1999; 99US-0162142P.

Query Match 100.0%; Score 76; DB 3; Length 101;  
Best Local Similarity 100.0%; Pred. No. 6.3e-06;  
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 DEAQSKRGILTLYP 15  
| | | | | | | | | | | | | | |  
Db 41 DEAQSKRGILTLYP 55

RESULT 15  
ABP42469

ID ABP42469 standard; protein; 105 AA.  
 AC ABP42469;  
 XX  
 DT 22-AUG-2002 (first entry)  
 XX  
 DE Human ovarian antigen HOCQ94, SEQ ID NO:3601.  
 XX  
 KW Human; ovarian antigen; ovary; ovarian; breast; cancer; tumour;  
 KW ovarian cancer; breast cancer; tumour; reproductive system disorder;  
 KW infertility; pregnancy disorder; anovulation; polycystic ovary syndrome;  
 KW PCOS; ovarian cyst; dysmenorrhoea; endocrine disorder; infection;  
 KW inflammatory condition; immune disorder; blood disorder;  
 KW cardiovascular disorder; respiratory disorder; neurological disorder;  
 KW gastrointestinal disorder; urinary system disorder; drug screening;  
 KW gene therapy; chromosome mapping; forensic analysis;  
 KW antibody preparation; cytostatic; immunomodulatory; neuroprotective;  
 KW antiinflammatory; gynaecological; reproductive.  
 XX  
 OS Homo sapiens.  
 XX  
 PN WO200200677-A1.  
 XX  
 PD 03-JAN-2002.  
 XX  
 PF 07-JUN-2001; 2001WO-US018569.  
 XX  
 PR 07-JUN-2000; 2000US-0209467P.  
 XX  
 PA (HUMA-) HUMAN GENOME SCI INC.  
 XX  
 PI Birse CE, Rosen CA;  
 XX  
 DR WPI; 2002-147878/19.  
 DR N-PSDB; ABQ55546.  
 XX  
 PT Isolated nucleic acid molecules encoding novel ovarian polypeptides,  
 PT useful in the prevention, treatment and diagnosis of cancer (e.g. ovarian  
 PT cancer), immune disorders, cardiovascular disorders and neurological  
 PT diseases.  
 XX  
 PS Claim 11; SEQ ID NO 3601; 2922pp; English.  
 XX  
 CC The invention relates to 2175 novel human ovarian antigens (ABP41054-  
 CC ABP43228) and to cDNAs encoding them (ABQ54131-ABQ56305), and also  
 CC encompasses polypeptides 90% identical and polynucleotides 95% identical  
 CC to the sequences of the invention. The invention additionally relates to  
 CC recombinant vectors and host cells comprising human ovarian antigen  
 CC polynucleotides, antibodies against human ovarian antigens, and the use  
 CC of ovarian antigen polynucleotides and polypeptides in diagnosing,  
 CC treating, prognosing or preventing various ovary and/or breast-related  
 CC disorders. Such conditions include ovarian cancer and breast cancer, and  
 CC metastatic tumours of ovarian or breast origin, reproductive system  
 CC disorders (e.g., infertility, disorders of pregnancy, anovulation,  
 CC polycystic ovary syndrome, ovarian cysts, and dysmenorrhoea), endocrine  
 CC disorders, infections (e.g., chlamydia, HIV, toxoplasmosis, and toxic  
 CC shock syndrome), inflammatory conditions (e.g., mastitis, oophoritis and  
 CC vaginitis), immune disorders (e.g., congenital and acquired  
 CC immunodeficiencies, autoimmune oophoritis, systemic lupus erythematosus),  
 CC blood-related disorders (e.g., anaemia), cardiovascular disorders,  
 CC respiratory disorders, neurological disorders, gastrointestinal disorders  
 CC and urinary system disorders. Ovarian antigen polypeptides and  
 CC polynucleotides may also be used in screening for compounds which  
 CC modulate ovarian antigen expression or activity. The polynucleotides may  
 CC further be used for gene therapy, chromosome mapping, in the  
 CC identification of individuals and in forensic analysis, and the  
 CC polypeptides may be used as food additives or to prepare antibodies  
 CC useful in disease diagnosis, drug targeting and phenotyping. The present  
 CC sequence represents a human ovarian antigen of the invention. Note: The  
 CC sequence data for this patent did not form part of the printed  
 CC specification, but was obtained in electronic format directly from WIPO  
 CC at ftp.wipo.int/pub/published\_pct\_sequences  
 XX

SQ Sequence 105 AA;  
 Query Match 100.0%; Score 76; DB 5; Length 105;  
 Best Local Similarity 100.0%; Pred. No. 6.5e-06;  
 Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 Qy 1 DEAQSKRGILTLKYP 15  
 Db 57 DEAQSKRGILTLKYP 71  
 |||||  
 Search completed: April 8, 2005, 10:50:48  
 Job time : 63.4 secs

**THIS PAGE BLANK (USPTO)**



GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: April 8, 2005, 09:58:56 ; Search time 62.4 Seconds  
(without alignments)  
92.971 Million cell updates/sec

Title: US-09-423-351C-8

Perfect score: 75

Sequence: 1 LASLSTFQQMWISKQ 15

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 2105692 seqs, 386760381 residues

Total number of hits satisfying chosen parameters: 2105692

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Listing first 45 summaries

Database : A\_Geneseq\_16Dec04:\*

- 1: geneseqp1980s:\*
- 2: geneseqp1990s:\*
- 3: geneseqp2000s:\*
- 4: geneseqp2001s:\*
- 5: geneseqp2002s:\*
- 6: geneseqp2003as:\*
- 7: geneseqp2003bs:\*
- 8: geneseqp2004s:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB ID	Description
1	75	100.0	15	2 AAW92534	Aaw92534 Beta-acti
2	75	100.0	48	8 ADP81223	Adp81223 Protein o
3	75	100.0	86	4 ABG08617	Abg08617 Novel hum
4	75	100.0	110	5 ABP42828	Abp42828 Human ova
5	75	100.0	129	4 AAM81819	Aam81819 Human hae
6	75	100.0	135	4 AAU32061	Aau32061 Novel hum
7	75	100.0	145	5 ABP43127	Abp43127 Human ova
8	75	100.0	147	4 AAM80875	Aam80875 Human hae
9	75	100.0	147	4 AAM81886	Aam81886 Human hae
10	75	100.0	147	4 AAM81332	Aam81332 Human hae
11	75	100.0	149	8 ABO84771	Abos84771 Human can
12	75	100.0	181	4 ABG11441	Abg11441 Novel hum
13	75	100.0	197	6 ADA55146	Ada55146 Human pro
14	75	100.0	250	4 ABG08618	Abg08618 Novel hum
15	75	100.0	274	6 ABU70549	Abu70549 Human adi
16	75	100.0	278	4 ABG15466	Abg15466 Novel hum
17	75	100.0	279	4 ABG11440	Abg11440 Novel hum
18	75	100.0	294	7 ADI63068	Adi63068 Human apo
19	75	100.0	332	7 ADI63011	Adi63011 Human apo
20	75	100.0	334	6 ABP98860	Abp98860 Human str
21	75	100.0	374	3 AAB12985	Aab12985 Human bet
22	75	100.0	375	1 AAP61532	Aap61532 Sequence
23	75	100.0	375	2 AAR22026	Aar22026 A. chryso
24	75	100.0	375	2 AAR22096	Aar22096 Actin. 3/
25	75	100.0	375	2 AAR49248	Aar49248 Actin. 3/

ALIGNMENTS

RESULT 1

AAW92534  
ID AAW92534 standard; peptide; 15 AA.

AC AAW92534;

XX 26-APR-1999 (first entry)

DE Beta-actin reference peptide substrate #8.

XX Peptide substrate; CCT; eukaryotic type II chaperonin complex; cyclin;  
KW binding agent; substrate-binding site; SBS; substrate folding; actin;  
KW tubulin; treatment; cancer; anticancer drug; viral infection; screening;  
KW reduced toxicity.

XX Synthetic.

XX WO9853322-A1.

XX 26-NOV-1998.

XX 22-MAY-1998; 98WO-GB001485.

XX 23-MAY-1997; 97GB-00010762.

XX (CANC-) INST CANCER RES ROYAL CANCER HOSPITAL.

XX Willison K, Hynes G, Liou AK;

XX WPI; 1999-070162/06.

PT Identifying specific binding agents for substrate binding site in CCT  
PT chaperonin complex, also new peptide binding agents and their mimetics,  
PT and peptides containing a specific CCT binding site, used for treating  
cancer.

PS Disclosure; Fig 10; 97pp; English.

CC This invention describes a method which uses the CCT (eukaryotic type II  
CC chaperonin) complex or part of it, for identifying a binding agent that  
CC can occupy a substrate-binding site (SBS) on the CCT complex. By binding  
CC to the CCT complex, the binding agents block an SBS so that biological  
CC activity of the CCT complex is affected, particularly its ability to fold  
CC substrates such as actin, tubulin and cyclin. The binding agents are  
CC useful for treatment of cancer, particularly when used in combination  
CC with an anticancer drug, or viral infections. Nucleic acid fragments are  
CC used to screen for agents, e.g. binding agents that modulate interaction  
CC between the CCT complex and a protein that is to be folded. The binding

Aar50328 Drug resi  
Aaw77100 P. chryso  
Aaw77101 A. chryso  
Aaw59176 P. calini  
Aab15017 Posttrans  
Aay94569 Human car  
Aab15016 Posttrans  
Aab577395 Human act  
Aab64271 Angiogene  
Abm04830 Rat cytop  
Adb85212 Rat actin  
Ade61174 Rat Prote  
Adf30525 Rat angio  
Adi62970 Human apo  
Adi62970 Human apo  
Adi63040 Human apo  
Adi13002 Human ste  
Adj78489 Actin pro  
Adl14103 Human ear  
Adn04506 Antipsori

CC agents may target cells that are actively synthesising tubulin etc.  
 CC (unlike known microtubule-stabilising agents that affect all cells), so  
 CC should have reduced toxicity for normal cells. AAW92527-W92541 are  
 CC peptide substrates used in the method of the invention  
 XX  
 SQ Sequence 15 AA;  
 Query Match 100.0%; Score 75; DB 2; Length 15;  
 Best Local Similarity 100.0%; Pred. No. 6.9e-06;  
 Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 LASLSTFOQMWISKQ 15  
 DB 1 LASLSTFOQMWISKQ 15  
 RESULT 2  
 ADP81223  
 ID ADP81223 standard; protein; 48 AA.  
 XX  
 AC ADP81223;  
 XX  
 DT 09-SEP-2004 (first entry)  
 XX  
 DE Protein of human ovarian specific gene, SEQ ID No 257.  
 XX  
 KW normal; neoplastic; ovarian; ovarian specific nucleic acid; OSNA;  
 KW metastatic; cancer; vaccine; cytostatic; human.  
 XX  
 OS Homo sapiens.  
 XX  
 PN WO2004053079-A2.  
 XX  
 PD 24-JUN-2004.  
 XX  
 PF 08-DEC-2003; 2003WO-US038855.  
 XX  
 PR 06-DEC-2002; 2002US-0431301P.  
 PR 06-DEC-2002; 2002US-0431321P.  
 PR 30-JUN-2003; 2003US-0484584P.  
 PR 07-NOV-2003; 2003US-0518607P.  
 XX  
 PA (DIAD-) DIADEXUS INC.  
 XX  
 PI Macina RA, Turner LR, Sun Y, Liu S, Chen H;  
 XX  
 DR WPI; 2004-468850/44.  
 DR N-PSDB; ADP81063.  
 XX  
 PT New ovarian specific nucleic acid molecules and polypeptides useful for  
 PT diagnosing, preventing or treating ovarian cancer, for producing  
 PT transgenic animals or cells, or for research purposes.  
 XX  
 PS Claim 12; SEQ ID NO 257; 754pp; English.  
 XX  
 CC The invention relates to novel isolated nucleic acid molecules and  
 CC polypeptides present in normal and neoplastic ovarian cells. These  
 CC comprise a nucleic acid sequence encoding any of the 167 amino acid  
 CC sequences (e.g. 438, 237 or 233 amino acids) fully defined in the  
 CC specification (SEQ. ID NOS: ADP81095 to ADP81261) and comprises any of  
 CC the 128 nucleotide sequences (e.g. 4798, 1494 or 1691 bp) fully defined  
 CC in the specification (SEQ. ID NOS: ADP80967 to ADP81094). The invention  
 CC further comprises a method for determining the presence of a ovarian  
 CC specific nucleic acid (OSNA) in a sample; a vector comprising the above  
 CC nucleic acid molecule; a host cell comprising the vector; a method for  
 CC producing a polypeptide encoded by the above nucleic acid molecule; a  
 CC polypeptide encoded by the nucleic acid molecule cited above; an antibody  
 CC or its fragment that specifically binds to the above polypeptide; a  
 CC method for determining the presence of an ovarian specific protein in a  
 CC sample; a method for diagnosing or monitoring the presence and metastases  
 CC of ovarian cancer in a patient; a kit for detecting a risk of cancer or  
 CC presence of cancer in a patient, the kit comprising a means for  
 CC determining the presence of the above nucleic acid molecule or

CC polypeptide; a method of treating a patient with ovarian cancer; and a  
 CC vaccine comprising the above polypeptide or nucleic acid encoding the  
 CC polypeptide. The isolated nucleic acid molecules and polypeptides have  
 CC cytostatic activity. The isolated polypeptides may be used to create a  
 CC vaccine. The isolated nucleic acid molecules and polypeptides can be used  
 CC for diagnosing or monitoring the presence and metastases of ovarian  
 CC cancer and treating ovarian cancer. This sequence represents the protein  
 CC of an ovarian specific gene of the invention.  
 XX  
 SQ Sequence 48 AA;  
 Query Match 100.0%; Score 75; DB 8; Length 48;  
 Best Local Similarity 100.0%; Pred. No. 2.2e-05;  
 Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 LASLSTFOQMWISKQ 15  
 DB 19 LASLSTFOQMWISKQ 33  
 RESULT 3  
 ABG08617  
 ID ABG08617 standard; protein; 86 AA.  
 XX  
 AC ABG08617;  
 XX  
 DT 13-FEB-2002 (first entry)  
 XX  
 DE Novel human diagnostic protein #8608.  
 XX  
 KW Human; chromosome mapping; gene mapping; gene therapy; forensic;  
 KW food supplement; medical imaging; diagnostic; genetic disorder.  
 XX  
 OS Homo sapiens.  
 XX  
 PN WO200175067-A2.  
 XX  
 PD 11-OCT-2001.  
 XX  
 PF 30-MAR-2001; 2001WO-US008631.  
 XX  
 PR 31-MAR-2000; 2000US-00540217.  
 PR 23-AUG-2000; 2000US-00649167.  
 XX  
 PA (HVSE-) HVSEQ INC.  
 XX  
 PI Drmanac RT, Liu C, Tang YT;  
 XX  
 DR WPI; 2001-639362/73.  
 DR N-PSDB; AAS72804.  
 XX  
 PT New isolated polynucleotide and encoded polypeptides, useful in  
 PT diagnostics, forensics, gene mapping, identification of mutations  
 PT responsible for genetic disorders or other traits and to assess  
 PT biodiversity.  
 XX  
 PS Claim 20; SEQ ID NO 38976; 103pp; English.  
 XX  
 CC The invention relates to isolated polynucleotide (I) and polypeptide (II)  
 CC sequences. (I) is useful as hybridisation probes, polymerase chain  
 CC reaction (PCR) primers, oligomers, and for chromosome and gene mapping,  
 CC and in recombinant production of (II). The polynucleotides are also used  
 CC in diagnostics as expressed sequence tags for identifying expressed  
 CC genes. (I) is useful in gene therapy techniques to restore normal  
 CC activity of (II) or to treat disease states involving (II). (II) is  
 CC useful for generating antibodies against it, detecting or quantitating a  
 CC polypeptide in tissue, as molecular weight markers and as a food  
 CC supplement. (II) and its binding partners are useful in medical imaging  
 CC of sites expressing (II). (I) and (II) are useful for treating disorders  
 CC involving aberrant protein expression or biological activity. The  
 CC polypeptide and polynucleotide sequences have applications in  
 CC diagnostics, forensics, gene mapping, identification of mutations  
 CC responsible for genetic disorders or other traits to assess biodiversity

CC and to produce other types of data and products dependent on DNA and  
 CC amino acid sequences. ABG00010-ABG30377 represent novel human diagnostic  
 CC amino acid sequences of the invention. Note: The sequence data for this  
 CC patent did not appear in the printed specification, but was obtained in  
 CC electronic format directly from WIPO at  
 CC ftp.wipo.int/pub/published\_pct\_sequences

XX SQ Sequence 86 AA;

Query Match 100.0%; Score 75; DB 4; Length 86;  
 Best Local Similarity 100.0%; Pred. No. 4e-05;  
 Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 LASLSTFQQMWISQ 15  
 |||||  
 Db 57 LASLSTFQQMWISQ 71

RESULT 4  
 ABP42928  
 ID ABP42928 standard; protein; 110 AA.

XX AC ABP42928;

XX DT 22-AUG-2002 (first entry)

XX DE Human ovarian antigen HPRS46, SEQ ID NO:4060.

XX Human; ovarian antigen; ovary; ovarian; breast; cancer; tumour;  
 KW ovarian cancer; breast cancer; tumour; reproductive system disorder;  
 KW infertility; pregnancy disorder; anovulation; polycystic ovary syndrome;  
 KW PCOS; ovarian cyst; dysmenorrhoea; endocrine disorder; infection;  
 KW inflammatory condition; immune disorder; blood disorder;  
 KW cardiovascular disorder; respiratory disorder; neurological disorder;  
 KW gastrointestinal disorder; urinary system disorder; drug screening;  
 KW gene therapy; chromosome mapping; forensic analysis;  
 KW antibody preparation; cytostatic; immunomodulatory; neuroprotective;  
 KW antiinflammatory; gynaecological; reproductive.

XX OS Homo sapiens.

XX FN WO200200677-A1.

XX PD 03-JAN-2002.

XX PF 07-JUN-2001; 2001WO-US018569.

XX PR 07-JUN-2000; 2000US-0209467P.

XX PA (HUMA-) HUMAN GENOME SCI INC.

XX PI Birse CE, Rosen CA;

XX DR WPI; 2002-147878/19.

XX DR N-PSDB; ABQ56005.

XX Isolated nucleic acid molecules encoding novel ovarian polypeptides,  
 PT useful in the prevention, treatment and diagnosis of cancer (e.g. ovarian  
 PT cancer), immune disorders, cardiovascular disorders and neurological  
 PT diseases.

XX PS Claim 11; SEQ ID NO 4060; 2922pp; English.

XX The invention relates to 2175 novel human ovarian antigens (ABP41054-  
 CC ABP43228) and to cDNAs encoding them (ABQ54131-ABQ56305), and also  
 CC encompasses polypeptides 90% identical and polynucleotides 95% identical  
 CC to the sequences of the invention. The invention additionally relates to  
 CC recombinant vectors and host cells comprising human ovarian antigen  
 CC polynucleotides, antibodies against human ovarian antigens, and the use  
 CC of ovarian antigen polynucleotides and polypeptides in diagnosing,  
 CC treating, prognosing or preventing various ovary and/or breast-related  
 CC disorders. Such conditions include ovarian cancer and breast cancer, and  
 CC metastatic tumours of ovarian or breast origin, reproductive system

CC disorders (e.g., infertility, disorders of pregnancy, anovulation,  
 CC polycystic ovary syndrome, ovarian cysts, and dysmenorrhoea), endocrine  
 CC disorders, infections (e.g., chlamydia, HIV, toxoplasmosis, and toxic  
 CC shock syndrome), inflammatory conditions (e.g., mastitis, oophoritis and  
 CC vaginitis), immune disorders (e.g., congenital and acquired  
 CC immunodeficiencies, autoimmune oophoritis, systemic lupus erythematosus),  
 CC blood-related disorders (e.g., anaemia), cardiovascular disorders,  
 CC respiratory disorders, neurological disorders, gastrointestinal disorders  
 CC and urinary system disorders. Ovarian antigen polypeptides and  
 CC polynucleotides may also be used in screening for compounds which  
 CC modulate ovarian antigen expression or activity. The polynucleotides may  
 CC further be used for gene therapy, chromosome mapping, in the  
 CC identification of individuals and in forensic analysis, and the  
 CC polypeptides may be used as food additives or to prepare antibodies  
 CC useful in disease diagnosis, drug targeting and phenotyping. The present  
 CC sequence represents a human ovarian antigen of the invention. Note: The  
 CC sequence data for this patent did not form part of the printed  
 CC specification, but was obtained in electronic format directly from WIPO  
 CC at ftp.wipo.int/pub/published\_pct\_sequences

XX SQ Sequence 110 AA;

Query Match 100.0%; Score 75; DB 5; Length 110;

Best Local Similarity 100.0%; Pred. No. 5.2e-05;  
 Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 LASLSTFQQMWISQ 15

Db 81 LASLSTFQQMWISQ 95

RESULT 5

AAM81819

ID AAM81819 standard; protein; 129 AA.

XX AC AAM81819;

XX DT 13-NOV-2001 (first entry)

XX DE Human haematological malignancy-related antigen #1517.

XX Human; cytostatic; vascular; gene therapy; vaccine; lymphoma;  
 KW haematological malignancy; antigen; chronic lymphocytic leukaemia;  
 KW follicular lymphoma; Hodgkin's lymphoma; non-Hodgkin's lymphoma.

XX OS Homo sapiens.

XX FN WO200164886-A2.

XX PD 07-SEP-2001.

XX PF 01-MAR-2001; 2001WO-US007272.

XX PR 01-MAR-2000; 2000US-0186126P.

XX PR 17-MAR-2000; 2000US-0190479P.

XX PR 27-APR-2000; 2000US-0200545P.

XX PR 28-APR-2000; 2000US-0200303P.

XX PR 28-APR-2000; 2000US-0200779P.

XX PR 01-MAY-2000; 2000US-0200999P.

XX PR 04-MAY-2000; 2000US-0202084P.

XX PR 22-MAY-2000; 2000US-0206201P.

XX PR 14-JUL-2000; 2000US-0218950P.

XX PR 03-AUG-2000; 2000US-0222903P.

XX PR 04-AUG-2000; 2000US-0223416P.

XX PR 07-AUG-2000; 2000US-0223378P.

XX PA (CORI-) CORIXA CORP.

XX PI Gaiger A, Algate PA, Mannion J;

XX DR WPI; 2001-514842/56.

XX Compositions and methods for the detection of hematological malignancies,

PT e.g. chronic lymphocytic leukemia, lymphoma, follicular lymphoma and  
 PT Hodgkin's and T/B cell non-Hodgkin's lymphoma.  
 XX  
 PS Claim 1; Page 1114; 1252pp; English.  
 XX  
 CC The present invention relates to compositions and methods for the  
 CC detection, diagnosis and therapy of haematological malignancies. The  
 CC present sequence is the protein sequence of a human haematological  
 CC malignancy related antigen. The methods of the present invention comprise  
 CC detecting the presence of haematological malignancy related antigen(s) in  
 CC a sample obtained from the patient (an increased level of the  
 CC polypeptide, compared to an unaffected individual, is indicative of an  
 CC increased risk). Haematological malignancies which can be treated using  
 CC the present invention are chronic lymphocytic leukaemia, lymphoma,  
 CC follicular lymphoma, Hodgkin's lymphoma, non-Hodgkin's lymphoma and T/B  
 CC cell non-Hodgkin's lymphoma  
 XX  
 SQ Sequence 129 AA;

Query Match 100.0%; Score 75; DB 4; Length 129;  
 Best Local Similarity 100.0%; Pred. No. 6.2e-05;  
 Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 LASLSTFQQMWISKQ 15  
 |||||  
 Db 100 LASLSTFQQMWISKQ 114

RESULT 6  
 AAU32061  
 ID AAU32061 standard; protein; 135 AA.  
 XX  
 AC AAU32061;  
 XX  
 DT 18-DEC-2001 (first entry)  
 XX  
 DE Novel human secreted protein #2552.  
 XX  
 KW Human; vaccination; gene therapy; nutritional supplement;  
 KW stem cell proliferation; haematopoiesis; nerve tissue regeneration;  
 KW immune suppression; immune stimulation; anti-inflammatory; leukaemia.  
 XX  
 OS Homo sapiens.  
 XX  
 XX WO200179449-A2.  
 XX  
 XX 25-OCT-2001.  
 XX  
 XX 16-APR-2001; 2001WO-US008656.  
 XX  
 XX 18-APR-2000; 2000US-00552929.  
 XX  
 XX 26-JAN-2001; 2001US-00770160.  
 XX  
 XX (HYSE-) HYSEQ INC.  
 XX  
 XX Tang YT, Liu C, Drmanac RT;  
 XX  
 XX WPI; 2001-611725/70.  
 XX  
 XX Nucleic acids encoding a range of human polypeptides, useful in genetic  
 XX vaccination, testing and therapy.  
 XX  
 XX Claim 20; Page 550; 765pp; English.

XX The invention relates to novel human secreted polypeptides. The  
 CC polypeptides and antibodies to the polypeptides are useful for  
 CC determining the presence of or predisposition to a disease associated  
 CC with altered levels of polypeptide. The polypeptides are also useful for  
 CC identifying agents (agonists and antagonists) that bind to them. Cells  
 CC expressing the proteins are useful for identifying a therapeutic agent  
 CC for use in treatment of a pathology related to aberrant expression or  
 CC physiological interactions of the polypeptide. Vectors comprising the  
 CC nucleic acids encoding the polypeptides and cells genetically engineered

CC to express them are also useful for producing the proteins. The proteins  
 CC are useful in genetic vaccination, testing and therapy, and can be used  
 CC as nutritional supplements. They may be used to increase stem cell  
 CC proliferation; to regulate haematopoiesis; and in bone, cartilage, tendon  
 CC stimulation; as anti-inflammatory agents; and in treatment of leukaemias.  
 CC AAU29510-AAU33304 represent the amino acid sequences of novel human  
 CC secreted proteins of the invention  
 XX  
 SQ Sequence 135 AA;

Query Match 100.0%; Score 75; DB 4; Length 135;  
 Best Local Similarity 100.0%; Pred. No. 6.5e-05;  
 Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 LASLSTFQQMWISKQ 15  
 |||||  
 Db 97 LASLSTFQQMWISKQ 111

RESULT 7  
 ABP43127  
 ID ABP43127 standard; protein; 145 AA.

XX  
 AC ABP43127;

XX  
 DT 22-AUG-2002 (first entry)

XX  
 DE Human ovarian antigen HVBD91, SEQ ID NO:4259.

XX Human; ovarian antigen; ovary; ovarian; breast; cancer; tumour;  
 KW ovarian cancer; breast cancer; tumour; reproductive system disorder;  
 KW infertility; pregnancy disorder; anovulation; polycystic ovary syndrome;  
 KW PCOS; ovarian cyst; dysmenorrhea; endocrine disorder; infection;  
 KW inflammatory condition; immune disorder; blood disorder;  
 KW cardiovascular disorder; respiratory disorder; neurological disorder;  
 KW gastrointestinal disorder; urinary system disorder; drug screening;  
 KW gene therapy; chromosome mapping; forensic analysis;  
 KW antibody preparation; cytostatic; immunomodulatory; neuroprotective;  
 KW antiinflammatory; gynaecological; reproductive.

XX  
 OS Homo sapiens.

XX  
 PN WO200200677-A1.

XX  
 PD 03-JAN-2002.

XX  
 PF 07-JUN-2001; 2001WO-US018569.

XX  
 PR 07-JUN-2000; 2000US-0209467P.

XX  
 PA (HUMA-) HUMAN GENOME SCI INC.

XX  
 PI Birse CE, Rosen CA;

XX  
 DR WPI; 2002-147878/19.

XX  
 DR N-PSDB; ABQ56204.

XX Isolated nucleic acid molecules encoding novel ovarian polypeptides,  
 PT useful in the prevention, treatment and diagnosis of cancer (e.g. ovarian  
 PT cancer), immune disorders, cardiovascular disorders and neurological  
 PT diseases.

XX  
 PS Claim 11; SEQ ID NO 4259; 2922pp; English.

XX The invention relates to 2175 novel human ovarian antigens (ABP41054-  
 CC ABP43228) and to cDNAs encoding them (ABQ54131-ABQ56305), and also  
 CC encompasses polypeptides 90% identical and polynucleotides 95% identical  
 CC to the sequences of the invention. The invention additionally relates to  
 CC recombinant vectors and host cells comprising human ovarian antigen  
 CC polynucleotides, antibodies against human ovarian antigens, and the use  
 CC of ovarian antigen polynucleotides and polypeptides in diagnosing,  
 CC treating, prognosing or preventing various ovary and/or breast-related

disorders. Such conditions include ovarian cancer and breast cancer, and metastatic tumours of ovarian or breast origin, reproductive system disorders (e.g., infertility, disorders of pregnancy, anovulation, polycystic ovary syndrome, ovarian cysts, and dysmenorrhoea), endocrine disorders, infections (e.g., chlamydia, HIV, toxoplasmosis, and toxic shock syndrome), inflammatory conditions (e.g., mastitis, oophoritis and vaginitis), immune disorders (e.g., congenital and acquired immunodeficiencies, autoimmune oophoritis, systemic lupus erythematosus), blood-related disorders (e.g., anaemia), cardiovascular disorders, respiratory disorders, neurological disorders, gastrointestinal disorders and urinary system disorders. Ovarian antigen polypeptides and modulate ovarian antigen expression or activity. The polynucleotides may further be used for gene therapy, chromosome mapping, in the identification of individuals and in forensic analysis, and the polypeptides may be used as food additives or to prepare antibodies useful in disease diagnosis, drug targeting and phenotyping. The present sequence represents a human ovarian antigen of the invention. Note: The sequence data for this patent did not form part of the printed specification, but was obtained in electronic format directly from WIPO at ftp.wipo.int/pub/published\_pct\_sequences

XX Sequence 145 AA;

Query Match 100.0%; Score 75; DB 5; Length 145;  
 Best Local Similarity 100.0%; Pred. No. 7e-05;  
 Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 LASLSTFOQMWSKQ 15  
 |||||  
 Db 116 LASLSTFOQMWSKQ 130

RESULT 8  
 AAM80875  
 ID AAM80875 standard; protein; 147 AA.

AC AAM80875;

DT 13-NOV-2001 (first entry)

DE Human haematological malignancy-related antigen #573.

KW Human; cytostatic; vascular; gene therapy; vaccine; lymphoma;  
 haematological malignancy; antigen; chronic lymphocytic leukaemia;  
 follicular lymphoma; Hodgkin's lymphoma; non-Hodgkin's lymphoma.

XX Homo sapiens.

OS WO200164886-A2.

PN 07-SEP-2001.

PD 01-MAR-2001; 2001WO-US007272.

PF 01-MAR-2000; 2000US-0186126P.

PR 17-MAR-2000; 2000US-0190479P.

PR 27-APR-2000; 2000US-0200545P.

PR 28-APR-2000; 2000US-0200303P.

PR 01-MAY-2000; 2000US-0200779P.

PR 04-MAY-2000; 2000US-0200999P.

PR 22-MAY-2000; 2000US-0202084P.

PR 14-JUL-2000; 2000US-0218950P.

PR 03-AUG-2000; 2000US-0222903P.

PR 04-AUG-2000; 2000US-0223416P.

PR 07-AUG-2000; 2000US-0223378P.

XX (CORI-) CORIXA CORP.

PI Gaiger A, Algate PA, Mannion J;

XX WPI; 2001-514842/56.

XX Compositions and methods for the detection of hematological malignancies,  
 PT e.g. chronic lymphocytic leukemia, lymphoma, follicular lymphoma and  
 PT Hodgkin's and T/B cell non-Hodgkin's lymphoma.

XX Claim 1; Page 743-744; 1252pp; English.

XX The present invention relates to compositions and methods for the  
 CC detection, diagnosis and therapy of haematological malignancies. The  
 CC present sequence is the protein sequence of a human haematological  
 CC malignancy related antigen. The methods of the present invention comprise  
 CC detecting the presence of haematological malignancy related antigen(s) in  
 CC a sample obtained from the patient (an increased level of the  
 CC polypeptide, compared to an unaffected individual, is indicative of an  
 CC increased risk). Haematological malignancies which can be treated using  
 CC the present invention are chronic lymphocytic leukaemia, lymphoma,  
 CC follicular lymphoma, Hodgkin's lymphoma, non-Hodgkin's lymphoma and T/B  
 CC cell non-Hodgkin's lymphoma

XX Sequence 147 AA;

Query Match 100.0%; Score 75; DB 4; Length 147;  
 Best Local Similarity 100.0%; Pred. No. 7.1e-05;  
 Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 LASLSTFOQMWSKQ 15  
 |||||  
 Db 118 LASLSTFOQMWSKQ 132

RESULT 9  
 AAM81886  
 ID AAM81886 standard; protein; 147 AA.

AC AAM81886;

DT 13-NOV-2001 (first entry)

DE Human haematological malignancy-related antigen #1584.

KW Human; cytostatic; vascular; gene therapy; vaccine; lymphoma;  
 haematological malignancy; antigen; chronic lymphocytic leukaemia;  
 follicular lymphoma; Hodgkin's lymphoma; non-Hodgkin's lymphoma.

XX Homo sapiens.

OS WO200164886-A2.

PN 07-SEP-2001.

PD 01-MAR-2001; 2001WO-US007272.

PF 01-MAR-2000; 2000US-0186126P.

PR 17-MAR-2000; 2000US-0190479P.

PR 27-APR-2000; 2000US-0200545P.

PR 28-APR-2000; 2000US-0200303P.

PR 01-MAY-2000; 2000US-0200779P.

PR 04-MAY-2000; 2000US-0202084P.

PR 14-JUL-2000; 2000US-0218950P.

PR 03-AUG-2000; 2000US-0222903P.

PR 04-AUG-2000; 2000US-0223416P.

PR 07-AUG-2000; 2000US-0223378P.

XX (CORI-) CORIXA CORP.

PI Gaiger A, Algate PA, Mannion J;

XX WPI; 2001-514842/56.

XX Compositions and methods for the detection of hematological malignancies,  
 PT e.g. chronic lymphocytic leukemia, lymphoma, follicular lymphoma and

PT Hodgkin's and T/B cell non-Hodgkin's lymphoma.  
 XX Claim 1; Page 1141; 1252pp; English.  
 PS  
 CC The present invention relates to compositions and methods for the  
 CC detection, diagnosis and therapy of haematological malignancies. The  
 CC present sequence is the protein sequence of a human haematological  
 CC malignancy related antigen. The methods of the present invention comprise  
 CC detecting the presence of haematological malignancy related antigen(s) in  
 CC a sample obtained from the patient (an increased level of the  
 CC polypeptide, compared to an unaffected individual, is indicative of an  
 CC increased risk). Haematological malignancies which can be treated using  
 CC the present invention are chronic lymphocytic leukaemia, lymphoma,  
 CC follicular lymphoma, Hodgkin's lymphoma, non-Hodgkin's lymphoma and T/B  
 CC cell non-Hodgkin's lymphoma  
 XX  
 SQ Sequence 147 AA;

Query Match 100.0%; Score 75; DB 4; Length 147;  
 Best Local Similarity 100.0%; Pred. No. 7.1e-05;  
 Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 LASLSTFQQMWISKQ 15  
 |||||  
 DB 118 LASLSTFQQMWISKQ 132

RESULT 10  
 AAM81332  
 ID AAM81332 standard; protein; 147 AA.  
 AC  
 XX AAM81332;  
 DT  
 DT 13-NOV-2001 (first entry)  
 XX  
 DE Human haematological malignancy-related antigen #1030.  
 DE  
 KW Human; cytostatic; vascular; gene therapy; vaccine; lymphoma;  
 KW haematological malignancy; antigen; chronic lymphocytic leukaemia;  
 KW follicular lymphoma; Hodgkin's lymphoma; non-Hodgkin's lymphoma.  
 XX  
 OS Homo sapiens.

XX WO200164886-A2.  
 PN  
 XX  
 PD 07-SEP-2001.

XX 01-MAR-2001; 2001WO-US007272.  
 PF

XX 01-MAR-2000; 2000US-0186126P.  
 PR 17-MAR-2000; 2000US-0190479P.  
 PR 27-APR-2000; 2000US-0200545P.  
 PR 28-APR-2000; 2000US-0200303P.  
 PR 28-APR-2000; 2000US-0200779P.  
 PR 01-MAY-2000; 2000US-0200999P.  
 PR 04-MAY-2000; 2000US-0202084P.  
 PR 14-MAY-2000; 2000US-0206201P.  
 PR 03-AUG-2000; 2000US-0218950P.  
 PR 04-AUG-2000; 2000US-0222903P.  
 PR 07-AUG-2000; 2000US-0223416P.  
 XX 07-AUG-2000; 2000US-0223378P.

XX (CORI-) CORIXA CORP.  
 PA

XX Gaiger A, Algate PA, Mannion J;  
 PI

XX WPI; 2001-514842/56.  
 DR

XX Compositions and methods for the detection of hematological malignancies,  
 PT e.g. chronic lymphocytic leukemia, lymphoma, follicular lymphoma and  
 PT Hodgkin's and T/B cell non-Hodgkin's lymphoma.  
 XX

PS Claim 1; Page 925; 1252pp; English.

XX  
 CC The present invention relates to compositions and methods for the  
 CC detection, diagnosis and therapy of haematological malignancies. The  
 CC present sequence is the protein sequence of a human haematological  
 CC malignancy related antigen. The methods of the present invention comprise  
 CC detecting the presence of haematological malignancy related antigen(s) in  
 CC a sample obtained from the patient (an increased level of the  
 CC polypeptide, compared to an unaffected individual, is indicative of an  
 CC increased risk). Haematological malignancies which can be treated using  
 CC the present invention are chronic lymphocytic leukaemia, lymphoma,  
 CC follicular lymphoma, Hodgkin's lymphoma, non-Hodgkin's lymphoma and T/B  
 CC cell non-Hodgkin's lymphoma  
 XX  
 SQ Sequence 147 AA;

Query Match 100.0%; Score 75; DB 4; Length 147;  
 Best Local Similarity 100.0%; Pred. No. 7.1e-05;  
 Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 LASLSTFQQMWISKQ 15  
 |||||  
 DB 118 LASLSTFQQMWISKQ 132

RESULT 11  
 ABO84771  
 ID ABO84771 standard; protein; 149 AA.  
 AC

XX ABO84771;  
 AC

XX 18-NOV-2004 (first entry)  
 DT

XX Human cancer-associated protein (CAP) HP07-010.  
 DT

DE Human; cancer-associated protein; CAP; cancer; cytostatic.  
 KW

XX Homo sapiens.  
 OS

XX WO2004058146-A2.  
 PN

XX 15-JUL-2004.  
 PD

XX 15-DEC-2003; 2003WO-US040081.  
 PF

XX 17-DEC-2002; 2002US-00322281.  
 PR

XX (SAGR-) SAGRES DISCOVERY INC.  
 XX

XX Morris DW, Malandro MS;  
 XX

XX WPI; 2004-499109/47.  
 XX N-PSDB; ABD31111.

XX Novel human cancer associated protein encoded within open reading frame  
 of cancer associated gene, useful as targets for diagnosing cancer.

XX Claim 18; SEQ ID NO 66; 182pp; English.

XX The invention relates to cancer-associated proteins (CAP) and the cancer-  
 associated (CA) nucleic acids encoding them. The invention also relates  
 to a method for treating cancers involving administering to a patient an  
 inhibitor of CAP, and a method of screening for anticancer activity in a  
 potential drug involving providing a cell that expresses a CA gene,  
 contacting a tissue sample derived from a cancer cell with an anticancer  
 drug candidate and monitoring the effect of the anticancer drug candidate  
 on expression of the CA gene. The CAP proteins are useful for detecting  
 cancer associated with expression of a CAP protein in a test cell sample  
 and for screening for a bioactive agent capable of modulating the  
 activity of a CAP protein. The CA nucleic acids are useful for diagnosing  
 cancer, involving determining the expression of a CA nucleic acid in a  
 tissue. This sequence represents a human CAP of the invention. Note: The  
 sequence data for this patent did not form part of the printed  
 specification, but was obtained in electronic format directly from WIPO

```
CC at ftp.wipo.int/pub/published_pct_sequences
XX
SQ Sequence 149 AA;
    Query Match      100.0%; Score 75; DB 8; Length 149;
    Best Local Similarity 100.0%; Pred. No. 7.2e-05;
    Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Oy 1 LASLSTFOQMWSKQ 15
    |||||
Db 120 LASLSTFOQMWSKQ 134
    |||||

RESULT 12
ABG11441
ID ABG11441 standard; protein; 181 AA.
XX
AC ABG11441;
XX
DT 18-FEB-2002 (first entry)
XX
DE Novel human diagnostic protein #11432.
XX
KW Human; chromosome mapping; gene mapping; gene therapy; forensic;
KW food supplement; medical imaging; diagnostic; genetic disorder.
XX
OS Homo sapiens.
XX
PN WO200175067-A2.
XX
PD 11-OCT-2001.
XX
PF 30-MAR-2001; 2001WO-US008631.
XX
PR 31-MAR-2000; 2000US-00540217.
XX
PR 23-AUG-2000; 2000US-00649167.
XX
PA (HYSE-) HYSEQ INC.
XX
PI Dmanac RT, Liu C, Tang YT;
XX
DR WPI; 2001-639362/73.
XX
DR N-PSDB; AAS75628.
XX
PT New isolated polynucleotide and encoded polypeptides, useful in
PT diagnostics, forensics, gene mapping, identification of mutations
PT responsible for genetic disorders or other traits and to assess
PT biodiversity.
XX
PS Claim 20; SEQ ID NO 41800; 103pp; English.
XX
CC The invention relates to isolated polynucleotide (I) and polypeptide (II)
CC sequences. (I) is useful as hybridisation probes, polymerase chain
CC reaction (PCR) primers, oligomers, and for chromosome and gene mapping,
CC and in recombinant production of (II). The polynucleotides are also used
CC in diagnostics as expressed sequence tags for identifying expressed
CC genes. (II) is useful in gene therapy techniques to restore normal
CC activity of (II) or to treat disease states involving (II). (II) is
CC useful for generating antibodies against it, detecting or quantitating a
CC polypeptide in tissue, as molecular weight markers and as a food
CC supplement. (II) and its binding partners are useful in medical imaging
CC of sites expressing (II). (I) and (II) are useful for treating disorders
CC involving aberrant protein expression or biological activity. The
CC polypeptide and polynucleotide sequences have applications in
CC diagnostics, forensics, gene mapping, identification of mutations
CC responsible for genetic disorders or other traits to assess biodiversity
CC and to produce other types of data and products dependent on DNA and
CC amino acid sequences. ABG00010-ABG30377 represent novel human diagnostic
CC amino acid sequences of the invention. Note: The sequence data for this
CC patent did not appear in the printed specification, but was obtained in
CC electronic format directly from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences
XX
```

```

SQ Sequence 181 AA;
    Query Match      100.0%; Score 75; DB 4; Length 181;
    Best Local Similarity 100.0%; Pred. No. 8.8e-05;
    Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Oy 1 LASLSTFOQMWSKQ 15
    |||||
Db 152 LASLSTFOQMWSKQ 166
    |||||

RESULT 13
ADA55146
ID ADA55146 standard; protein; 197 AA.
XX
AC ADA55146;
XX
DT 20-NOV-2003 (first entry)
XX
DE Human protein, SEQ ID 2714.
XX
KW Cytostatic; Anti-inflammatory; Osteopathic; Neuroprotective; Nootropic;
KW Gene therapy; human; secretory protein; membrane proteins; cancer;
KW inflammatory disease; osteoporosis; neurological disease.
XX
OS Homo sapiens.
XX
PN EP1293569-A2.
XX
PD 19-MAR-2003.
XX
PF 21-MAR-2002; 2002EP-00006586.
XX
PR 14-SEP-2001; 2001JP-00328381.
XX
PR 24-JAN-2002; 2002US-0350435P.
XX
PA (HELI-) HELIX RES INST.
PA (REAS-) RES ASSOC BIOTECHNOLOGY.
XX
PI Isogai T, Sugiyama T, Otsuki T, Wakamatsu A, Sato H, Ishii S;
PI Yamamoto J, Isono Y, Hio Y, Otsuka K, Nagai K, Irie R, Tamechika I;
PI Seki N, Yoshikawa T, Otsuka M, Nagahari K, Masuho Y;
XX
DR WPI; 2003-395539/38.
XX
DR N-PSDB; ADA53507.
XX
PT New polynucleotides encoding full-length polypeptides, e.g. secretory
PT and/or membrane proteins, useful for developing medicines for diseases in
PT which the gene is involved, or as target molecules for gene therapy.
XX
PS Claim 14; SEQ ID NO 2714; 205pp; English.
XX
CC The present invention relates to novel human secretory or membrane
CC proteins (ADA54072-ADA55710) and their coding sequences (ADA52433-
CC ADA54071). The coding sequences are useful in the gene therapy of
CC diseases caused by abnormalities of the proteins, e.g. cancer,
CC inflammatory diseases, osteoporosis or neurological disease.
XX
SQ Sequence 197 AA;
    Query Match      100.0%; Score 75; DB 6; Length 197;
    Best Local Similarity 100.0%; Pred. No. 9.6e-05;
    Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Oy 1 LASLSTFOQMWSKQ 15
    |||||
Db 168 LASLSTFOQMWSKQ 182
    |||||

RESULT 14
ABG08618
ID ABG08618 standard; protein; 250 AA.
XX
```

AC	ABG08618;	AC
XX		
DT	13-FEB-2002 (first entry)	
XX		
DE	Novel human diagnostic protein #8609.	
XX		
KW	Human; chromosome mapping; gene mapping; gene therapy; forensic;	
KW	food supplement; medical imaging; diagnostic; genetic disorder.	
XX		
OS	Homo sapiens...	
XX		
PN	WO200175067-A2.	
XX		
PD	11-OCT-2001.	
XX		
XX	30-MAR-2001; 2001WO-US008631.	
XX		
PR	31-MAR-2000; 2000US-00540217.	
PR	23-AUG-2000; 2000US-00649167.	
XX		
PA	(HYSE-) HYSEQ INC.	
XX		
PI	Drmanac RT, Liu C, Tang YT;	
XX		
DR	WPI; 2001-639362/73.	
DR	N-PSDB; AAS72805.	
XX		
PT	New isolated polynucleotide and encoded polypeptides, useful in	
PT	diagnostics, forensics, gene mapping, identification of mutations	
PT	responsible for genetic disorders or other traits and to assess	
PT	biodiversity.	
XX		
PS	Claim 20; SEQ ID NO 38977; 103pp; English.	
XX		
CC	The invention relates to isolated polynucleotide (I) and polypeptide (II)	
CC	sequences. (I) is useful as hybridisation probes, polymerase chain	
CC	reaction (PCR) primers, oligomers, and for chromosome and gene mapping,	
CC	and in recombinant production of (II). The polynucleotides are also used	
CC	in diagnostics as expressed sequence tags for identifying expressed	
CC	genes. (I) is useful in gene therapy techniques to restore normal	
CC	activity of (II) or to treat disease states involving (II). (II) is	
CC	useful for generating antibodies against it, detecting or quantitating a	
CC	polypeptide in tissue, as molecular weight markers and as a food	
CC	supplement. (II) and its binding partners are useful in medical imaging	
CC	of sites expressing (II). (I) and (II) are useful for treating disorders	
CC	involving aberrant protein expression or biological activity. The	
CC	polypeptide and polynucleotide sequences have applications in	
CC	diagnostics, forensics, gene mapping, identification of mutations	
CC	responsible for genetic disorders or other traits to assess biodiversity	
CC	and to produce other types of data and products dependent on DNA and	
CC	amino acid sequences. ABG00010-ABG30377 represent novel human diagnostic	
CC	amino acid sequences of the invention. Note: The sequence data for this	
CC	patent did not appear in the printed specification, but was obtained in	
CC	electronic format directly from WIPO at	
CC	ftp.wipo.int/pub/published_pct_sequences	
XX		
XX	Sequence 250 AA;	
XX		
XX	Query Match 100.0%; Score 75; DB 4; Length 250;	
XX	Best Local Similarity 100.0%; Pred. No. 0.00012;	
XX	Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0	
QY	1 LASTSTFQQMWSKQ 15	
DB	220 LASTSTFQQMWSKQ 234	
RESULT 15		
ABU70549		
ID	ABU70549 standard; protein; 274 AA.	
XX		
AC	ABU70549;	
XX		

DT	10-JUN-2003 (first entry)	
XX		
DE	Human adipocyte Selected Interacting domain, SID, #180.	
XX		
KW	Human; prey; adipocyte; SID; selected interacting domain; anorectic;	
KW	antiadipetic; protein-protein interaction; diabetes;	
KW	yeast 2-hybrid assay; metabolic disorder; obesity.	
XX		
OS	Homo sapiens.	
XX		
XX	W0200286122-A2.	
PN		
XX	31-OCT-2002.	
PD		
XX		
PF	14-MAR-2002; 2002WO-EP003769.	
XX		
PR	14-MAR-2001; 2001US-0275734P.	
XX		
PA	(HYBR-) HYBRIGENICS.	
XX		
PI	Legrain P, Daviet L;	
XX		
DR	WPI; 2003-103412/09.	
DR	N-PSDB; ACA57093.	
XX		
PT	New complex between two interacting proteins in adipocyte cells, useful	
PT	for identifying selected interacting domains that modulate protein	
PT	interactions, or for preventing or treating metabolic disorders such as	
PT	obesity or diabetes.	
XX		
PS	Claim 6; Page 167; 382pp; English.	
XX		
CC	The invention relates to a complex between two interacting proteins in	
CC	adipocyte cells, given in the specification. The proteins are identified	
CC	by selecting a bait protein from a known adipocyte marker and then	
CC	performing a yeast 2-hybrid selection to isolate prey proteins encoded by	
CC	members of an adipocyte cDNA library. The proteins are designated SID	
CC	(RTM) (selected interacting domains) proteins. Also included are a	
CC	polynucleotide encoding a polypeptide in the adipocyte cells, a	
CC	recombinant host cell expressing at least one of the interacting	
CC	polypeptides of the complex, selecting a modulating compound in adipocyte	
CC	cells, a SID (RTM) polypeptide comprising any of the 738 amino acid	
CC	sequences given in the specification (including its fragment or variant),	
CC	a SID (RTM) polynucleotide comprising any of the 738 nucleotide sequences	
CC	given in the specification (including its fragment or variant), a vector	
CC	comprising the SID (RTM) polynucleotide, a recombinant host cell	
CC	comprising the vector, a protein chip comprising the polypeptides and a	
CC	record comprising all or part of the data, listed in the specification.	
CC	The complex, polypeptides, polynucleotides and compounds are useful for	
CC	preventing or treating metabolic disorders such as obesity or diabetes.	
CC	The polynucleotides are useful as probes or primers. The complex is	
CC	particularly useful for identifying selected interacting domains (SID	
CC	(RTM)) for screening drugs that modulate the protein interaction, thus	
CC	exhibiting the therapeutic effect. The present sequence represents a SID	
CC	(prey) protein of the invention	
XX		
SQ	Sequence 274 AA;	
	Query Match 100.0%; Score 75; DB 6; Length 274;	
	Best Local Similarity 100.0%; Pred. No. 0.00014;	
	Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0	
QY	1 LASLSTFOQMWSKQ 15	
Db	245 LASLSTFOQMWSKQ 259	

Search completed: April 8, 2005, 10:50:47  
Job time : 63.4 secs



GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

# OM protein - protein search, using sw model

Run on: April 8, 2005, 09:58:57; Search time 17.5333 Seconds  
(without alignments)  
63.863 Million cell updates/sec

Title: US-09-423-351C-8

Perfect score: 75

Sequence: 1 LASLSTFQQMWSKQ 15

Scoring table: BLOSUM62

Gapop 10.0, Gapext 0.5

Searched: 513545 seqs, 74649064 residues

Total number of hits satisfying chosen parameters: 513545

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

## Database :

Issued Patents\_AA\*

- 1: /cgn2\_6/ptodata/1/iaa/5A\_COMB.pep.\*
- 2: /cgn2\_6/ptodata/1/iaa/5B\_COMB.pep.\*
- 3: /cgn2\_6/ptodata/1/iaa/6A\_COMB.pep.\*
- 4: /cgn2\_6/ptodata/1/iaa/6B\_COMB.pep.\*
- 5: /cgn2\_6/ptodata/1/iaa/PCTUS\_COMB.pep.\*
- 6: /cgn2\_6/ptodata/1/iaa/backfiles.pep.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	75	100.0	239	4	US-09-949-016-10409
2	75	100.0	374	3	US-08-609-236-6
3	75	100.0	375	3	US-09-106-217-16
4	75	100.0	375	3	US-09-171-337A-7
5	75	100.0	375	4	US-09-631-022-7
6	75	100.0	375	4	US-09-976-594-731
7	75	100.0	377	3	US-09-106-217-2
8	75	100.0	377	4	US-09-919-172-33
9	75	100.0	377	4	US-09-917-254-53
10	75	100.0	377	4	US-09-248-796A-14109
11	75	100.0	386	4	US-09-949-016-7721
12	75	100.0	399	4	US-09-949-016-9424
13	75	100.0	404	4	US-09-949-016-11313
14	70	93.3	23	4	US-09-700-436-4
15	70	93.3	375	2	US-08-494-151-14
16	70	93.3	375	3	US-09-171-337A-8
17	70	93.3	375	4	US-09-631-022-8
18	70	93.3	376	1	US-08-588-113-2
19	70	93.3	376	4	US-09-949-016-6100
20	70	93.3	402	4	US-09-949-016-10757
21	65	86.7	47	3	US-09-306-446C-7
22	65	86.7	374	3	US-09-306-446C-2
23	64	85.3	164	4	US-09-270-767-33348
24	64	85.3	164	4	US-09-270-767-48565
25	58	77.3	362	4	US-09-949-016-7725
26	58	77.3	376	4	US-09-538-092-1109
27	58	77.3	376	4	US-09-538-092-1110

28	58	77.3	376	4	US-09-949-016-6656	Sequence 6656, Ap
29	58	77.3	376	4	US-09-949-016-8452	Sequence 8452, Ap
30	51	68.0	336	4	US-09-248-796A-14108	Sequence 14108, A
31	47	62.7	497	4	US-09-248-796A-18965	Sequence 18965, A
32	46	61.3	384	4	US-09-538-092-395	Sequence 395, App
33	41	54.7	313	4	US-09-198-452A-794	Sequence 794, App
34	41	54.7	333	4	US-09-438-185A-746	Sequence 746, App
35	40	53.3	1805	4	US-09-949-016-8246	Sequence 8246, App
36	39	52.0	161	4	US-09-270-767-40150	Sequence 40150, A
37	39	52.0	161	4	US-09-270-767-55366	Sequence 55366, A
38	39	52.0	419	4	US-09-328-352-6451	Sequence 6451, Ap
39	39	52.0	460	4	US-09-489-039A-13505	Sequence 13505, A
40	38	50.7	73	4	US-09-621-976-6616	Sequence 6616, Ap
41	38	50.7	84	4	US-09-328-352-4655	Sequence 4655, Ap
42	38	50.7	300	4	US-09-252-991A-26038	Sequence 26038, A
43	38	50.7	347	4	US-09-543-681A-6024	Sequence 6024, Ap
44	38	50.7	603	4	US-09-396-149-8	Sequence 8, Appli
45	38	50.7	940	4	US-09-328-352-8165	Sequence 8165, Ap

## ALIGNMENTS

### RESULT 1

US-09-949-016-10409  
; Sequence 10409, Application US/09949016  
; Patent No. 6812339  
; GENERAL INFORMATION:  
; APPLICANT: VENTER, J. Craig et al.  
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF  
; FILE REFERENCE: CLO01307  
; CURRENT APPLICATION NUMBER: US/09/949,016  
; CURRENT FILING DATE: 2000-04-14  
; PRIOR APPLICATION NUMBER: 60/241,755  
; PRIOR FILING DATE: 2000-10-20  
; PRIOR APPLICATION NUMBER: 60/237,768  
; PRIOR FILING DATE: 2000-10-03  
; PRIOR APPLICATION NUMBER: 60/231,498  
; PRIOR FILING DATE: 2000-09-08  
; NUMBER OF SEQ ID NOS: 207012  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 10409  
; LENGTH: 239  
; TYPE: PRT  
; ORGANISM: Human  
US-09-949-016-10409

Query Match 100.0%; Score 75; DB 4; Length 239;  
Best Local Similarity 100.0%; Pred. No. 2.6e-05;  
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 LASLSTFQQMWSKQ 15  
Db 210 LASLSTFQQMWSKQ 224

### RESULT 2

US-08-609-236-6  
; Sequence 6, Application US/08609236  
; Patent No. 6087398  
; GENERAL INFORMATION:  
; APPLICANT: Steven R. Goodman  
; TITLE OF INVENTION: No. 6087398el Sickle Cell Anemia Treatment  
; NUMBER OF SEQUENCES: 6  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: McGregor & Adler, LLP  
; STREET: 8011 Candle Lane  
; CITY: Houston  
; STATE: Texas  
; COUNTRY: USA  
; ZIP: 77071  
; COMPUTER READABLE FORM:

```

; MEDIUM TYPE: Floppy disk
; COMPUTER: Apple Macintosh
; OPERATING SYSTEM: Macintosh
; SOFTWARE: Microsoft Word for Macintosh
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/609,236
; FILING DATE: March 1, 1996
; CLASSIFICATION: 514
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 60/002,288
; FILING DATE: August 14, 1995
; CLASSIFICATION: 514
; ATTORNEY/AGENT INFORMATION:
; NAME: Benjamin Aaron Adler, Ph.D.
; REGISTRATION NUMBER: 35,423
; REFERENCE/DOCKET NUMBER: D5807
; TELEPHONE: 713-777-2321
; TELEFAX: 713-777-6908
; INFORMATION FOR SEQ ID NO: 6:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 374
; TYPE: Amino acid
; STRANDEDNESS:
; TOPOLOGY: Linear
; MOLECULE TYPE: Protein
; DESCRIPTION: Protein
; HYPOTHETICAL: NO
; ANTI-SENSE: No
; FRAGMENT TYPE:
; ORIGINAL SOURCE:
; STRAIN:
; INDIVIDUAL ISOLATE:
; DEVELOPMENTAL STAGE:
; TISSUE TYPE:
; CELL TYPE:
; CELL LINE:
;
US-08-609-236-6

```

Query Match 100.0%; Score 75; DB 3; Length 374;

Best Local Similarity 100.0%; Pred. No. 4.3e-05;

Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

Qy 1 LASLSTFQQMWISKQ 15
Db 345 LASLSTFQQMWISKQ 359

```

#### RESULT 3

US-09-106-217-16

Sequence 16, Application US/09106217

Patent No. 6063576

GENERAL INFORMATION:

APPLICANT: Keating, Mark T.

APPLICANT: Olson, Timothy M.

TITLE OF INVENTION: Actin Mutations in Dilated

TITLE OF INVENTION: Cardiomyopathy, a Heritable Form of Heart Failure

NUMBER OF SEQUENCES: 18

CORRESPONDENCE ADDRESS:

ADDRESSEE: Rothwell, Figg, Ernst & Kurz, P.C.

STREET: 555 Thirteenth Street, N.W., Suite 701 East

CITY: Washington

STATE: DC

COUNTRY: U.S.A.

ZIP: 20004

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: PatentIn Release #1.0, Version #1.30

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/09/106,217

```

; FILING DATE:
; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
; NAME: Saxe, Stephen A.
; REGISTRATION NUMBER: 38,609
; REFERENCE/DOCKET NUMBER: 2323-125
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 202-783-6040
; TELEFAX: 202-783-6031
; INFORMATION FOR SEQ ID NO: 16:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 375 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
;
US-09-106-217-16

```

Query Match 100.0%; Score 75; DB 3; Length 375;

Best Local Similarity 100.0%; Pred. No. 4.3e-05;

Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

Qy 1 LASLSTFQQMWISKQ 15
Db 346 LASLSTFQQMWISKQ 360

```

#### RESULT 4

US-09-171-337A-7

Sequence 7, Application US/09171337A

Patent No. 6300095

GENERAL INFORMATION:

APPLICANT: BARREDO FUENTE, Jose Luis

RODRIGUEZ SAIZ, Marta

COLLADOS DE LA VIEJA, Alfonso J.

MORENO VALLE, Migueu Angel

SALTO MALDONADO, Francisco

DIEZ GARCIA, Bruno

TITLE OF INVENTION: PROMOTERS OF THE GENES GLUTAMATE  
DESHYDROGENASE, -N-ACETYLHEXOSAMINIDASE  
AND -ACTIN AND THEIR USE IN FILAMENTOUS  
FUNGI EXPRESSION, SECRETION AND ANTISENSE

NUMBER OF SEQUENCES: 20

CORRESPONDENCE ADDRESS:

ADDRESSEE: LADAS & PARRY

STREET: 26 WEST 61 STREET

CITY: NEW YORK

STATE: NY

COUNTRY: USA

ZIP: 10023

COMPUTER READABLE FORM:

MEDIUM TYPE: 3-1/4" Disk 1.44MB

COMPUTER: IBM PC compatible

OPERATING SYSTEM: Microsoft Windows for Workgroups 3.11

SOFTWARE: Wordperfect 8 for Windows

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/09/171,337A

FILING DATE: 14-May-1999

CLASSIFICATION: 536

PRIOR APPLICATION DATA:

APPLICATION NUMBER: PCT/ES98/00056

FILING DATE: 5-MAR-1998

APPLICATION NUMBER: ES9700482

FILING DATE: 5-MAR-1997

ATTORNEY/AGENT INFORMATION:

NAME: MASS, Clifford J.

REGISTRATION NUMBER: 30,086

(C) REF./DOCKET NO.: U-011948-3

TELECOMMUNICATION INFORMATION:

TELEPHONE: 233288

INFORMATION FOR SEQ ID NO: 7

SEQUENCE CHARACTERISTICS:

LENGTH: 375 amino acids

TYPE: amino acids

STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
ORIGINAL SOURCE: Penicillium chrysogenum  
FEATURE:  
OTHER INFORMATION: amino acid sequence of the -actin  
protein with a molecular weight of  
41760 Da.  
SEQUENCE DESCRIPTION: SEQ ID NO: 7  
US-09-171-337A-7

Query Match 100.0%; Score 75; DB 3; Length 375;  
Best Local Similarity 100.0%; Pred. No. 4.3e-05;  
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 LASLSTFOQMWSKQ 15  
Db 346 LASLSTFOQMWSKQ 360

## RESULT 5

US-09-631-022-7  
; Sequence 7, Application US/09631022  
; Patent No. 6558921  
; GENERAL INFORMATION:  
; APPLICANT: BARREDO FUENTE, Jose Luis  
; RODRIGUEZ SAIZ, Marta  
; COLLADOS DE LA VIEJA, Alfonso J.  
; MORENO VALLE, Migueu Angel  
; SALTO MALDONADO, Francisco  
; DIEZ GARCIA, Bruno  
; TITLE OF INVENTION: PROMOTERS OF THE GENES GLUTAMATE  
DESHYDROGENASE, -N-ACETYLHEXOSAMINIDASE  
AND -ACTIN AND THEIR USE IN FILAMENTOUS  
FUNGI EXPRESSION, SECRETION AND ANTISENSE  
NUMBER OF SEQUENCES: 20  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: LADAS & PARRY  
STREET: 26 WEST 61 STREET  
CITY: NEW YORK  
STATE: NY  
COUNTRY: USA  
ZIP: 10023  
COMPUTER READABLE FORM:  
MEDIUM TYPE: 3-1/4" Disk 1.44MB  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: Microsoft Windows for Workgroups 3.11  
SOFTWARE: WordPerfect 8 for Windows  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/631,022  
FILING DATE: 02-Aug-2000  
CLASSIFICATION DATA:  
PRIORITY APPLICATION DATA:  
APPLICATION NUMBER: 09/171,337  
FILING DATE: 14-MAY-1999  
APPLICATION NUMBER: PCT/ES98/00056  
FILING DATE: 5-MAR-1998  
APPLICATION NUMBER: ES9700482  
FILING DATE: 5-MAR-1997  
ATTORNEY/AGENT INFORMATION:  
NAME: MASS, Clifford J.  
REGISTRATION NUMBER: 30,086  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 233288  
INFORMATION FOR SEQ ID NO: 7  
SEQUENCE CHARACTERISTICS:  
LENGTH: 375 amino acids  
TYPE: amino acids  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: peptide

ORIGINAL SOURCE:  
ORGANISM: Penicillium chrysogenum  
FEATURE:  
OTHER INFORMATION: amino acid sequence of the -actin  
protein with a molecular weight of  
41760 Da.  
SEQUENCE DESCRIPTION: SEQ ID NO: 7  
US-09-631-022-7

Query Match 100.0%; Score 75; DB 4; Length 375;  
Best Local Similarity 100.0%; Pred. No. 4.3e-05;  
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 LASLSTFOQMWSKQ 15  
Db 346 LASLSTFOQMWSKQ 360

## RESULT 6

US-09-976-594-731  
; Sequence 731, Application US/09976594  
; Patent No. 6673549  
; GENERAL INFORMATION:  
; APPLICANT: Buchbinder, Jenny  
; TITLE OF INVENTION: GENES EXPRESSED IN C3A LIVER CELL CULTURES TREATED WITH STEROIDS  
; FILE REFERENCE: PA-0041 US  
; CURRENT APPLICATION NUMBER: US/09/976,594  
; CURRENT FILING DATE: 2001-10-12  
; PRIOR APPLICATION NUMBER: 60/240,409  
; PRIOR FILING DATE: 2000-10-12  
; NUMBER OF SEQ ID NOS: 1143  
; SOFTWARE: PERL Program  
; SEQ ID NO 731  
; LENGTH: 375  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
; FEATURE:  
; NAME/KEY: misc feature  
; OTHER INFORMATION: Incyte ID No. 6673549 1837317CD1  
US-09-976-594-731

Query Match 100.0%; Score 75; DB 4; Length 375;  
Best Local Similarity 100.0%; Pred. No. 4.3e-05;  
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 LASLSTFOQMWSKQ 15  
Db 346 LASLSTFOQMWSKQ 360

## RESULT 7

US-09-106-217-2  
; Sequence 2, Application US/09106217  
; Patent No. 6063576  
; GENERAL INFORMATION:  
; APPLICANT: Keating, Mark T.  
; APPLICANT: Olson, Timothy M.  
; TITLE OF INVENTION: Actin Mutations in Dilated  
; TITLE OF INVENTION: Cardiomyopathy, a Heritable Form of Heart Failure  
; NUMBER OF SEQUENCES: 18  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Rothwell, Figg, Ernst & Kurz, P.C.  
; STREET: 555 Thirteenth Street, N.W., Suite 701 East  
; STREET: Tower  
; CITY: Washington  
; STATE: DC  
; COUNTRY: U.S.A.  
; ZIP: 20004  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS

; SOFTWARE: PatentIn Release #1.0, Version #1.30  
; CURRENT APPLICATION NUMBER: US 09/106,217  
; FILING DATE: 2000-07-28  
; APPLICATION NUMBER: US 09/106,217  
; CLASSIFICATION:  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Saxe, Stephen A.  
; REGISTRATION NUMBER: 38,609  
; REFERENCE/DOCKET NUMBER: 2323-125  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 202-783-6040  
; TELEFAX: 202-783-6031  
; INFORMATION FOR SEQ ID NO: 2:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 377 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
US-09-106-217-2

Query Match 100.0%; Score 75; DB 3; Length 377;  
Best Local Similarity 100.0%; Pred. No. 4.3e-05;  
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 LASLSTFQQMWISKQ 15  
|||  
Db 348 LASLSTFQQMWISKQ 362

## RESULT 8

US-09-919-172-33  
; Sequence 33, Application US/09919172  
; Patent No. 6673545  
; GENERAL INFORMATION:  
; APPLICANT: Paris, Mary

; APPLICANT: Turner, Christopher M.  
; TITLE OF INVENTION: PROSTATE CANCER MARKERS  
; FILE REFERENCE: PA-0036 US  
; CURRENT APPLICATION NUMBER: US/09/919,172  
; CURRENT FILING DATE: 2001-07-30  
; PRIOR APPLICATION NUMBER: 60/222,469  
; PRIOR FILING DATE: 2000-07-28  
; NUMBER OF SEQ ID NOS: 102  
; SOFTWARE: PERL Program  
; SEQ ID NO 33  
; LENGTH: 377  
; TYPE: PRT  
; ORGANISM: Homo sapiens

; NAME/KEY: misc feature  
; OTHER INFORMATION: Incyte ID No. 6673545 1709118CD1  
US-09-919-172-33

Query Match 100.0%; Score 75; DB 4; Length 377;  
Best Local Similarity 100.0%; Pred. No. 4.3e-05;  
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 LASLSTFQQMWISKQ 15  
|||  
Db 348 LASLSTFQQMWISKQ 362

## RESULT 9

US-09-917-254-53  
; Sequence 53, Application US/09917254  
; Patent No. 6703204  
; GENERAL INFORMATION:  
; APPLICANT: Mutter, George

; APPLICANT: Baak, Jan  
; TITLE OF INVENTION: Prognostic Classification of Breast Cancer  
; FILE REFERENCE: B0801/7224(JRV)  
; CURRENT APPLICATION NUMBER: US/09/917,254  
; CURRENT FILING DATE: 2001-07-27

; PRIOR APPLICATION NUMBER: US 60/222,093  
; PRIOR FILING DATE: 2000-07-28  
; NUMBER OF SEQ ID NOS: 102  
; SOFTWARE: PatentIn version 3.0  
; SEQ ID NO 53  
; LENGTH: 377  
; TYPE: PRT  
; ORGANISM: Homo Sapiens  
US-09-917-254-53

Query Match 100.0%; Score 75; DB 4; Length 377;  
Best Local Similarity 100.0%; Pred. No. 4.3e-05;  
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 LASLSTFQQMWISKQ 15  
|||  
Db 348 LASLSTFQQMWISKQ 362

## RESULT 10

US-09-248-796A-14109  
; Sequence 14109, Application US/09248796A  
; Patent No. 6747137  
; GENERAL INFORMATION:  
; APPLICANT: Keith Weinstock et al

; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO CANDIDA ALBICAN;  
; FILE REFERENCE: 107196.132  
; CURRENT APPLICATION NUMBER: US/09/248,796A  
; CURRENT FILING DATE: 1999-02-12  
; PRIOR APPLICATION NUMBER: US 60/074,725  
; PRIOR FILING DATE: 1998-02-13  
; PRIOR APPLICATION NUMBER: US 60/096,409  
; PRIOR FILING DATE: 1998-08-13  
; NUMBER OF SEQ ID NOS: 28208  
; SEQ ID NO 14109  
; LENGTH: 377  
; TYPE: PRT  
; ORGANISM: Candida albicans

US-09-248-796A-14109

Query Match 100.0%; Score 75; DB 4; Length 377;  
Best Local Similarity 100.0%; Pred. No. 4.3e-05;  
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 LASLSTFQQMWISKQ 15  
|||  
Db 348 LASLSTFQQMWISKQ 362

## RESULT 11

US-09-949-016-7721  
; Sequence 7721, Application US/09949016  
; Patent No. 6812339  
; GENERAL INFORMATION:  
; APPLICANT: VENTER, J. Craig et al

; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED  
; FILE REFERENCE: CL001307  
; CURRENT APPLICATION NUMBER: US/09/949,016  
; CURRENT FILING DATE: 2000-04-14  
; PRIOR APPLICATION NUMBER: 60/241,755  
; PRIOR FILING DATE: 2000-10-20  
; PRIOR APPLICATION NUMBER: 60/237,768  
; PRIOR FILING DATE: 2000-10-03  
; PRIOR APPLICATION NUMBER: 60/231,498  
; PRIOR FILING DATE: 2000-09-08  
; NUMBER OF SEQ ID NOS: 207012  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 7721  
; LENGTH: 386  
; TYPE: PRT  
; ORGANISM: Human

## US-09-949-016-7721

Query Match 100.0%; Score 75; DB 4; Length 386;  
Best Local Similarity 100.0%; Pred. No. 4.5e-05;  
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 LASLSTFOQMWSKQ 15  
| | | | | | | | | | | | | | | | | |  
Db 357 LASLSTFOQMWSKQ 371

## RESULT 12

US-09-949-016-9424  
; Sequence 9424, Application US/09949016  
; Patent No. 6812339  
; GENERAL INFORMATION:  
; APPLICANT: VENTER, J. Craig et al.  
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED  
; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF  
; FILE REFERENCE: CL001307  
; CURRENT APPLICATION NUMBER: US/09/949,016  
; PRIOR FILING DATE: 2000-04-14  
; PRIOR APPLICATION NUMBER: 60/241,755  
; PRIOR FILING DATE: 2000-10-20  
; PRIOR APPLICATION NUMBER: 60/237,768  
; PRIOR FILING DATE: 2000-10-03  
; PRIOR APPLICATION NUMBER: 60/231,498  
; PRIOR FILING DATE: 2000-09-08  
; NUMBER OF SEQ ID NOS: 207012  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 9424  
; LENGTH: 399  
; TYPE: PRT  
; ORGANISM: Human  
US-09-949-016-9424

Query Match 100.0%; Score 75; DB 4; Length 399;  
Best Local Similarity 100.0%; Pred. No. 4.6e-05;  
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 LASLSTFOQMWSKQ 15  
| | | | | | | | | | | | | | | | | |  
Db 370 LASLSTFOQMWSKQ 384

## RESULT 13

US-09-949-016-11313  
; Sequence 11313, Application US/09949016  
; Patent No. 6812339  
; GENERAL INFORMATION:  
; APPLICANT: VENTER, J. Craig et al.  
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED  
; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF  
; FILE REFERENCE: CL001307  
; CURRENT APPLICATION NUMBER: US/09/949,016  
; PRIOR FILING DATE: 2000-04-14  
; PRIOR APPLICATION NUMBER: 60/241,755  
; PRIOR FILING DATE: 2000-10-20  
; PRIOR APPLICATION NUMBER: 60/237,768  
; PRIOR FILING DATE: 2000-10-03  
; PRIOR APPLICATION NUMBER: 60/231,498  
; PRIOR FILING DATE: 2000-09-08  
; NUMBER OF SEQ ID NOS: 207012  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 11313  
; LENGTH: 404  
; TYPE: PRT  
; ORGANISM: Human  
US-09-949-016-11313

Query Match 100.0%; Score 75; DB 4; Length 404;  
Best Local Similarity 100.0%; Pred. No. 4.7e-05;  
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 LASLSTFOQMWSKQ 15  
| | | | | | | | | | | | | | | | | |  
Db 375 LASLSTFOQMWSKQ 389

## RESULT 14

US-09-700-436-4  
; Sequence 4, Application US/09700436  
; Patent No. 6482802  
; GENERAL INFORMATION:  
; APPLICANT: Hu, Guo-Fu  
; APPLICANT: Vallee, Burt L.  
; TITLE OF INVENTION: USE OF NEOMYCIN FOR TREATING  
; ANGIOGENESIS RELATED DISEASES  
; FILE REFERENCE: 9457-008-999  
; CURRENT APPLICATION NUMBER: US/09/700,436  
; CURRENT FILING DATE: 2000-11-09  
; PRIOR APPLICATION NUMBER: PCT/US99/10269  
; PRIOR FILING DATE: 1999-05-11  
; PRIOR APPLICATION NUMBER: 60/084,921  
; PRIOR FILING DATE: 1998-05-11  
; NUMBER OF SEQ ID NOS: 5  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 4  
; LENGTH: 23  
; TYPE: PRT  
; ORGANISM: Homo Sapiens  
US-09-700-436-4

Query Match 93.3%; Score 70; DB 4; Length 23;  
Best Local Similarity 100.0%; Pred. No. 1.5e-05;  
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 LASLSTFOQMWSK 14  
| | | | | | | | | | | | | | | | | |  
Db 10 LASLSTFOQMWSK 23

## RESULT 15

US-08-494-151-14  
; Sequence 14, Application US/08494151  
; Patent No. 5840528  
; GENERAL INFORMATION:  
; APPLICANT: Van Ooyen, Albert Johannes Joseph  
; TITLE OF INVENTION: Transformation of Phaffia rhodozyma  
; NUMBER OF SEQUENCES: 14  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Morrison & Foerster  
; STREET: 2000 Pennsylvania Ave. N.W., Suite 5500  
; CITY: Washington, D.C.  
; COUNTRY: USA  
; ZIP: 20006-1812  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/494,151  
; FILING DATE: 23-JUN-1995  
; CLASSIFICATION: 435  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Murashige, Kate H.  
; REGISTRATION NUMBER: 29,959  
; REFERENCE/DOCKET NUMBER: 24615-20039.01  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (202) 887-1500  
; TELEFAX: (202) 887-0763  
; TELEX: 90-4030  
; INFORMATION FOR SEQ ID NO: 14:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 375 amino acids

; TYPE: amino acid  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
US-08-494-151-14

Query Match 93.3%; Score 70; DB 2; Length 375;  
Best Local Similarity 93.3%; Pred. No. 0.00031;  
Matches 14; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 LASLSTFQQMWSKQ 15  
Db 346 LASLSTFQSMWSKQ 360

Search completed: April 8, 2005, 12:07:37  
Job time : 18.5333 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: April 8, 2005, 10:53:18 ; Search time 42.6667 Seconds  
(without alignments)  
116.718 Million cell updates/sec

Title: US-09-423-351C-8

Perfect score: 75

Sequence: 1 LASLSTFOQMWISKQ 15

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1418010 seqs, 331997259 residues

Total number of hits satisfying chosen parameters: 1418010

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications AA:\*

- 1: /cgn2\_6/ptodata/1/pubpaa/US07\_PUBCOMB.pep.\*
- 2: /cgn2\_6/ptodata/1/pubpaa/PCT\_NEW PUB.pep.\*
- 3: /cgn2\_6/ptodata/1/pubpaa/US06\_NEW PUB.pep.\*
- 4: /cgn2\_6/ptodata/1/pubpaa/US06\_PUBCOMB.pep.\*
- 5: /cgn2\_6/ptodata/1/pubpaa/US07\_NEW PUB.pep.\*
- 6: /cgn2\_6/ptodata/1/pubpaa/PCTUS\_PUBCOMB.pep.\*
- 7: /cgn2\_6/ptodata/1/pubpaa/US08\_NEW PUB.pep.\*
- 8: /cgn2\_6/ptodata/1/pubpaa/US08\_PUBCOMB.pep.\*
- 9: /cgn2\_6/ptodata/1/pubpaa/US09A\_PUBCOMB.pep.\*
- 10: /cgn2\_6/ptodata/1/pubpaa/US09B\_PUBCOMB.pep.\*
- 11: /cgn2\_6/ptodata/1/pubpaa/US09C\_PUBCOMB.pep.\*
- 12: /cgn2\_6/ptodata/1/pubpaa/US09\_NEW PUB.pep.\*
- 13: /cgn2\_6/ptodata/1/pubpaa/US10A\_PUBCOMB.pep.\*
- 14: /cgn2\_6/ptodata/1/pubpaa/US10B\_PUBCOMB.pep.\*
- 15: /cgn2\_6/ptodata/1/pubpaa/US10C\_PUBCOMB.pep.\*
- 16: /cgn2\_6/ptodata/1/pubpaa/US10D\_PUBCOMB.pep.\*
- 17: /cgn2\_6/ptodata/1/pubpaa/US10\_NEW PUB.pep.\*
- 18: /cgn2\_6/ptodata/1/pubpaa/US11\_NEW PUB.pep.\*
- 19: /cgn2\_6/ptodata/1/pubpaa/US60\_NEW PUB.pep.\*
- 20: /cgn2\_6/ptodata/1/pubpaa/US60\_PUBCOMB.pep.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	75	100.0	47	15 US-10-341-434-139	Sequence 139, App
2	75	100.0	74	16 US-10-767-701-33598	Sequence 33598, A
3	75	100.0	81	15 US-10-424-599-157586	Sequence 157586
4	75	100.0	95	16 US-10-767-701-55472	Sequence 55472, A
5	75	100.0	110	15 US-10-284-049-4060	Sequence 4060, Ap
6	75	100.0	129	9 US-09-796-692-2183	Sequence 2183, Ap
7	75	100.0	129	14 US-10-040-862-2183	Sequence 2183, Ap
8	75	100.0	129	15 US-10-057-475B-2183	Sequence 2183, Ap
9	75	100.0	129	15 US-10-154-884B-2183	Sequence 2183, Ap
10	75	100.0	129	16 US-10-764-324-2183	Sequence 2183, Ap
11	75	100.0	133	15 US-10-424-599-143033	Sequence 143033
12	75	100.0	145	15 US-10-264-049-4259	Sequence 4259, Ap
13	75	100.0	147	9 US-09-796-692-1239	Sequence 1239, Ap

14	75	100.0	147	9 US-09-796-692-1696	Sequence 1696, Ap
15	75	100.0	147	9 US-09-796-692-2250	Sequence 2250, Ap
16	75	100.0	147	14 US-10-040-862-1239	Sequence 1239, Ap
17	75	100.0	147	14 US-10-040-862-1696	Sequence 1696, Ap
18	75	100.0	147	14 US-10-040-862-2250	Sequence 2250, Ap
19	75	100.0	147	15 US-10-057-475B-1239	Sequence 1239, Ap
20	75	100.0	147	15 US-10-057-475B-1696	Sequence 1696, Ap
21	75	100.0	147	15 US-10-057-475B-2250	Sequence 2250, Ap
22	75	100.0	147	15 US-10-154-884B-1239	Sequence 1239, Ap
23	75	100.0	147	15 US-10-154-884B-1696	Sequence 1696, Ap
24	75	100.0	147	15 US-10-154-884B-2250	Sequence 2250, Ap
25	75	100.0	147	16 US-10-322-281-66	Sequence 66, Appl
26	75	100.0	147	16 US-10-764-324-1239	Sequence 1239, Ap
27	75	100.0	147	16 US-10-764-324-1696	Sequence 1696, Ap
28	75	100.0	147	16 US-10-764-324-2250	Sequence 2250, Ap
29	75	100.0	197	15 US-10-094-749-2714	Sequence 2714, Ap
30	75	100.0	371	16 US-10-322-281-68	Sequence 68, Appl
31	75	100.0	375	14 US-10-205-194-93	Sequence 93, Appl
32	75	100.0	375	14 US-10-316-253-88	Sequence 88, Appl
33	75	100.0	375	15 US-10-369-493-2436	Sequence 2436, Ap
34	75	100.0	375	15 US-10-369-493-5927	Sequence 5927, Ap
35	75	100.0	375	15 US-10-205-331-94	Sequence 94, Appl
36	75	100.0	375	15 US-10-260-708-82	Sequence 82, Appl
37	75	100.0	377	9 US-09-919-172-33	Sequence 33, Appl
38	75	100.0	377	15 US-10-236-031B-64	Sequence 64, Appl
39	75	100.0	378	15 US-10-369-493-12372	Sequence 12372, A
40	75	100.0	393	15 US-10-092-900A-310	Sequence 310, App
41	75	100.0	398	16 US-10-322-281-63	Sequence 63, Appl
42	75	100.0	413	9 US-09-925-301-1436	Sequence 1436, Ap
43	75	100.0	448	15 US-10-369-493-4105	Sequence 4105, Ap
44	75	100.0	883	15 US-10-112-944-715	Sequence 715, App
45	72	96.0	38	16 US-10-767-701-47237	Sequence 47237, A

#### ALIGNMENTS

RESULT 1  
US-10-341-434-139  
; Sequence 139, Application US/10341434  
; Publication No. US20030215835A1  
; GENERAL INFORMATION:  
; APPLICANT: Origene Technologies  
; TITLE OF INVENTION: Differentially Regulated Prostate Cancer Genes  
; FILE REFERENCE: 9U 204 205 R1  
; CURRENT APPLICATION NUMBER: US/10/341,434  
; CURRENT FILING DATE: 2003-07-18  
; PRIOR APPLICATION NUMBER: US 60/348,164  
; PRIOR FILING DATE: 2002-01-15  
; PRIOR APPLICATION NUMBER: US 60/348,119  
; PRIOR FILING DATE: 2002-01-15  
; NUMBER OF SEQ ID NOS: 238  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 139  
; LENGTH: 47  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-10-341-434-139

Query Match 100.0%; Score 75; DB 15; Length 47;  
Best Local Similarity 100.0%; Pred. No. 2.8e-05;  
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 LASLSTFOQMWISKQ 15

Db 18 LASLSTFOQMWISKQ 32

#### RESULT 2

US-10-767-701-33598  
; Sequence 33598, Application US/10767701  
; Publication No. US20040172684A1  
; GENERAL INFORMATION:

; APPLICANT: Kovalic, David K.  
; APPLICANT: Zhou, Yihua  
; APPLICANT: Cao, Yongwei  
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With  
; TITLE OF INVENTION: Plants and Uses Thereof For Plant Improvement  
; FILE REFERENCE: 38-21(53535)B  
; CURRENT APPLICATION NUMBER: US/10/767,701  
; CURRENT FILING DATE: 2004-01-29  
; NUMBER OF SEQ ID NOS: 63128  
; SEQ ID NO 33598  
; LENGTH: 74  
; TYPE: PRT  
; ORGANISM: Sorghum bicolor  
; FEATURE:  
; OTHER INFORMATION: Clone ID: SORBI-28MAY03-C33956\_1.pep  
US-10-767-701-33598

Query Match 100.0%; Score 75; DB 16; Length 74;  
Best Local Similarity 100.0%; Pred. No. 4.4e-05;  
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 LASLSTFQQMWSKQ 15  
||| ||||| ||||| |||||  
Db 45 LASLSTFQQMWSKQ 59

RESULT 3  
US-10-424-599-157586  
; Sequence 157586, Application US/10424599  
; Publication No. US20040031072A1  
; GENERAL INFORMATION:  
; APPLICANT: La Rosa Thomas J  
; APPLICANT: Kovalic David K  
; APPLICANT: Zhou Yihua  
; APPLICANT: Cao Yongwei  
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With  
; TITLE OF INVENTION: Plants and Uses Thereof For Plant Improvement  
; FILE REFERENCE: 38-21(53223)B  
; CURRENT APPLICATION NUMBER: US/10/424,599  
; CURRENT FILING DATE: 2003-04-28  
; NUMBER OF SEQ ID NOS: 285684  
; SEQ ID NO 157586  
; LENGTH: 81  
; TYPE: PRT  
; ORGANISM: Glycine max  
; FEATURE:  
; OTHER INFORMATION: Clone ID: PAT\_MRT3847\_11331C.1.pep  
US-10-424-599-157586

Query Match 100.0%; Score 75; DB 15; Length 81;  
Best Local Similarity 100.0%; Pred. No. 4.8e-05;  
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 LASLSTFQQMWSKQ 15  
||| ||||| ||||| |||||  
Db 52 LASLSTFQQMWSKQ 66

RESULT 4  
US-10-767-701-55472  
; Sequence 55472, Application US/10767701  
; Publication No. US20040172684A1  
; GENERAL INFORMATION:  
; APPLICANT: Zhou, Yihua  
; APPLICANT: Kovalic, David K.  
; APPLICANT: Cao, Yongwei  
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With  
; TITLE OF INVENTION: Plants and Uses Thereof For Plant Improvement  
; FILE REFERENCE: 38-21(53535)B  
; CURRENT APPLICATION NUMBER: US/10/767,701  
; CURRENT FILING DATE: 2004-01-29  
; NUMBER OF SEQ ID NOS: 63128  
; SEQ ID NO 55472

; LENGTH: 95  
; TYPE: PRT  
; ORGANISM: Sorghum bicolor  
; FEATURE:  
; OTHER INFORMATION: Clone ID: 30163357.pep  
US-10-767-701-55472

Query Match 100.0%; Score 75; DB 16; Length 95;  
Best Local Similarity 100.0%; Pred. No. 5.6e-05;  
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 LASLSTFQQMWSKQ 15  
||| ||||| ||||| |||||  
Db 66 LASLSTFQQMWSKQ 80

RESULT 5  
US-10-264-049-4060  
; Sequence 4060, Application US/10264049  
; Publication No. US20040005579A1  
; GENERAL INFORMATION:  
; APPLICANT: Birse et al.  
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies  
; FILE REFERENCE: PA133P1  
; CURRENT APPLICATION NUMBER: US/10/264,049  
; CURRENT FILING DATE: 2002-10-04  
; PRIOR APPLICATION NUMBER: PCT/US01/18569  
; PRIOR FILING DATE: 2001-06-07  
; PRIOR APPLICATION NUMBER: US 60/209,467  
; PRIOR FILING DATE: 2000-06-07  
; NUMBER OF SEQ ID NOS: 4360  
; SOFTWARE: PatentIn Ver. 3.1  
; SEQ ID NO 4060  
; LENGTH: 110  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-10-264-049-4060

Query Match 100.0%; Score 75; DB 15; Length 110;  
Best Local Similarity 100.0%; Pred. No. 6.4e-05;  
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 LASLSTFQQMWSKQ 15  
||| ||||| ||||| |||||  
Db 81 LASLSTFQQMWSKQ 95

RESULT 6  
US-09-796-692-2183  
; Sequence 2183, Application US/09796692  
; Publication No. US20020198362A1  
; GENERAL INFORMATION:  
; APPLICANT: Gaiger, Alexander  
; APPLICANT: Algate, Paul A.  
; APPLICANT: Mannion, Jane  
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE DETECTION, DIAGNOSIS AND THERAPY  
; TITLE OF INVENTION: HEMATOLOGICAL MALIGNANCIES  
; FILE REFERENCE: 2077,001200  
; CURRENT APPLICATION NUMBER: US/09/796,692  
; CURRENT FILING DATE: 2001-03-01  
; PRIOR APPLICATION NUMBER: 60/186,126  
; PRIOR FILING DATE: 2000-03-01  
; PRIOR APPLICATION NUMBER: 60/190,479  
; PRIOR FILING DATE: 2000-03-17  
; PRIOR APPLICATION NUMBER: 60/200,545  
; PRIOR FILING DATE: 2000-04-27  
; PRIOR APPLICATION NUMBER: 60/200,303  
; PRIOR FILING DATE: 2000-04-28  
; PRIOR APPLICATION NUMBER: 60/200,779  
; PRIOR FILING DATE: 2000-04-28  
; PRIOR APPLICATION NUMBER: 60/200,999  
; PRIOR FILING DATE: 2000-05-01  
; PRIOR APPLICATION NUMBER: 60/202,084



```
; PRIOR FILING DATE: 2000-05-04
; PRIOR APPLICATION NUMBER: 60/206,201
; PRIOR FILING DATE: 2000-05-22
; PRIOR APPLICATION NUMBER: 60/218,950
; PRIOR FILING DATE: 2000-07-14
; PRIOR APPLICATION NUMBER: 60/222,903
; PRIOR FILING DATE: 2000-08-03
; PRIOR APPLICATION NUMBER: 60/223,416
; PRIOR FILING DATE: 2000-08-04
; PRIOR APPLICATION NUMBER: 60/223,378
; PRIOR FILING DATE: 2000-08-07
; NUMBER OF SEQ ID NOS: 9597
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 2183
; LENGTH: 129
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: variant
; LOCATION: (1)...(129)
; OTHER INFORMATION: Xaa = Any amino acid
US-09-796-692-2183
```

```
Query Match 100.0%; Score 75; DB 9; Length 129;
Best Local Similarity 100.0%; Pred. No. 7.5e-05;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
Qy 1 LASLSTFOQMWSKQ 15
Db 100 LASLSTFOQMWSKQ 114
```

## RESULT 7

```
US-10-040-862-2183
; Sequence 2183, Application US/10040862
; Publication No. US20030078396A1
; GENERAL INFORMATION:
; APPLICANT: Gaiger, Alexander
; APPLICANT: Algate, Paul A.
; APPLICANT: Mannion, Jane
; APPLICANT: Retter, Marc
; APPLICANT: Corixa Corporation
; TITLE OF INVENTION: Compositions and Methods for the Detection, Diagnosis and Therapy
; FILE REFERENCE: 014058-013520US
; CURRENT APPLICATION NUMBER: US/10/040,862
; CURRENT FILING DATE: 2001-11-06
; PRIOR APPLICATION NUMBER: US 60/186,126
; PRIOR FILING DATE: 2000-03-01
; PRIOR APPLICATION NUMBER: US 60/190,479
; PRIOR FILING DATE: 2000-03-17
; PRIOR APPLICATION NUMBER: US 60/200,545
; PRIOR FILING DATE: 2000-04-27
; PRIOR APPLICATION NUMBER: US 60/200,303
; PRIOR FILING DATE: 2000-04-28
; PRIOR APPLICATION NUMBER: US 60/200,779
; PRIOR FILING DATE: 2000-05-01
; PRIOR APPLICATION NUMBER: US 60/202,084
; PRIOR FILING DATE: 2000-05-04
; PRIOR APPLICATION NUMBER: US 60/206,201
; PRIOR FILING DATE: 2000-05-22
; PRIOR APPLICATION NUMBER: US 60/200,999
; PRIOR FILING DATE: 2000-07-14
; PRIOR APPLICATION NUMBER: US 60/218,950
; PRIOR FILING DATE: 2000-07-14
; PRIOR APPLICATION NUMBER: US 60/222,903
; PRIOR FILING DATE: 2000-08-03
; PRIOR APPLICATION NUMBER: US 60/223,416
; PRIOR FILING DATE: 2000-08-04
; PRIOR APPLICATION NUMBER: US 60/223,378
; PRIOR FILING DATE: 2000-08-07
; PRIOR APPLICATION NUMBER: US 09/796,692
; PRIOR FILING DATE: 2001-03-01
```

```
; NUMBER OF SEQ ID NOS: 10467
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 2183
; LENGTH: 129
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: variant
; LOCATION: (1)...(129)
; OTHER INFORMATION: Xaa = Any amino acid
US-10-040-862-2183
```

```
Query Match 100.0%; Score 75; DB 14; Length 129;
Best Local Similarity 100.0%; Pred. No. 7.5e-05;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
Qy 1 LASLSTFOQMWSKQ 15
Db 100 LASLSTFOQMWSKQ 114
```

## RESULT 8

```
US-10-057-475B-2183
; Sequence 2183, Application US/10057475B
; Publication No. US20040002068A1
; GENERAL INFORMATION:
; APPLICANT: Gaiger, Alexander
; APPLICANT: Algate, Paul A.
; APPLICANT: Mannion, Jane
; APPLICANT: Clapper, Jonathan David
; APPLICANT: Wang, Aijun
; APPLICANT: Ordonez, Nadia
; APPLICANT: Carter, Lauren
; APPLICANT: McNeill, Patricia Dianne
; APPLICANT: Corixa Corporation
; TITLE OF INVENTION: Compositions and Methods for the Detection, Diagnosis and Therapy
; FILE REFERENCE: 014058-014402US
; CURRENT APPLICATION NUMBER: US/10/057,475B
; CURRENT FILING DATE: 2002-01-22
; PRIOR APPLICATION NUMBER: US 60/186,126
; PRIOR FILING DATE: 2000-03-01
; PRIOR APPLICATION NUMBER: US 60/190,479
; PRIOR FILING DATE: 2000-03-17
; PRIOR APPLICATION NUMBER: US 60/200,545
; PRIOR FILING DATE: 2000-04-27
; PRIOR APPLICATION NUMBER: US 60/200,303
; PRIOR FILING DATE: 2000-04-28
; PRIOR APPLICATION NUMBER: US 60/200,779
; PRIOR FILING DATE: 2000-04-28
; PRIOR APPLICATION NUMBER: US 60/200,999
; PRIOR FILING DATE: 2000-05-01
; PRIOR APPLICATION NUMBER: US 60/202,084
; PRIOR FILING DATE: 2000-05-04
; PRIOR APPLICATION NUMBER: US 60/206,201
; PRIOR FILING DATE: 2000-05-22
; PRIOR APPLICATION NUMBER: US 60/218,950
; PRIOR FILING DATE: 2000-07-14
; PRIOR APPLICATION NUMBER: US 60/222,903
; PRIOR FILING DATE: 2000-08-03
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 10979
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 2183
; LENGTH: 129
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: variant
; LOCATION: (1)...(129)
; OTHER INFORMATION: Xaa = Any amino acid
US-10-057-475B-2183
```

```
Query Match      100.0%; Score 75; DB 15; Length 129;
Best Local Similarity 100.0%; Pred. No. 7.5e-05;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 LASLSTFQQMWISKQ 15
Db 100 LASLSTFQQMWISKQ 114

RESULT 9
US-10-154-884B-2183
; Sequence 2183, Application US/10154884B
; Publication No. US20040005561A1
; GENERAL INFORMATION:
; APPLICANT: Gaiger, Alexander
; APPLICANT: Algate, Paul A.
; APPLICANT: Mannion, Jane
; APPLICANT: Retter, Marc W.
; APPLICANT: Corixa Corporation
; TITLE OF INVENTION: Compositions and Methods for the Detection, Diagnosis and Therapy
; FILE REFERENCE: 014058-013521US
; CURRENT APPLICATION NUMBER: US/10/154,884B
; CURRENT FILING DATE: 2002-05-23
; PRIOR APPLICATION NUMBER: US 60/186,126
; PRIOR FILING DATE: 2000-03-01
; PRIOR APPLICATION NUMBER: US 60/190,479
; PRIOR FILING DATE: 2000-03-17
; PRIOR APPLICATION NUMBER: US 60/200,545
; PRIOR FILING DATE: 2000-04-28
; PRIOR APPLICATION NUMBER: US 60/200,779
; PRIOR FILING DATE: 2000-04-28
; PRIOR APPLICATION NUMBER: US 60/200,999
; PRIOR FILING DATE: 2000-05-01
; PRIOR APPLICATION NUMBER: US 60/202,084
; PRIOR FILING DATE: 2000-05-04
; PRIOR APPLICATION NUMBER: US 60/206,201
; PRIOR FILING DATE: 2000-05-22
; PRIOR APPLICATION NUMBER: US 60/218,950
; PRIOR FILING DATE: 2000-07-14
; PRIOR APPLICATION NUMBER: US 60/222,903
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 11290
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 2183
; LENGTH: 129
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: variant
; LOCATION: (1)...(129)
; OTHER INFORMATION: Xaa = Any amino acid
US-10-154-884B-2183

Query Match      100.0%; Score 75; DB 15; Length 129;
Best Local Similarity 100.0%; Pred. No. 7.5e-05;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 LASLSTFQQMWISKQ 15
Db 100 LASLSTFQQMWISKQ 114

RESULT 10
US-10-764-324-2183
; Sequence 2183, Application US/10764324
; Publication No. US20040175739A1
; GENERAL INFORMATION:
; APPLICANT: Gaiger, Alexander
; APPLICANT: Algate, Paul A.
```

```
; APPLICANT: Mannion, Jane
; APPLICANT: Retter, Marc
; APPLICANT: Corixa Corporation
; TITLE OF INVENTION: Compositions and Methods for the Detection, Diagnosis and Therapy
; FILE REFERENCE: 014058-013520US
; CURRENT APPLICATION NUMBER: US/10/764,324
; CURRENT FILING DATE: 2004-01-23
; PRIOR APPLICATION NUMBER: US/10/040,862
; PRIOR FILING DATE: 2001-11-06
; PRIOR APPLICATION NUMBER: US 60/186,126
; PRIOR FILING DATE: 2000-03-01
; PRIOR APPLICATION NUMBER: US 60/190,479
; PRIOR FILING DATE: 2000-03-17
; PRIOR APPLICATION NUMBER: US 60/200,545
; PRIOR FILING DATE: 2000-04-27
; PRIOR APPLICATION NUMBER: US 60/200,303
; PRIOR FILING DATE: 2000-04-28
; PRIOR APPLICATION NUMBER: US 60/200,779
; PRIOR FILING DATE: 2000-04-28
; PRIOR APPLICATION NUMBER: US 60/200,999
; PRIOR FILING DATE: 2000-05-01
; PRIOR APPLICATION NUMBER: US 60/202,084
; PRIOR FILING DATE: 2000-05-04
; PRIOR APPLICATION NUMBER: US 60/206,201
; PRIOR FILING DATE: 2000-05-22
; PRIOR APPLICATION NUMBER: US 60/218,950
; PRIOR FILING DATE: 2000-07-14
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 10467
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 2183
; LENGTH: 129
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: variant
; LOCATION: (1)...(129)
; OTHER INFORMATION: Xaa = Any amino acid
US-10-764-324-2183

Query Match      100.0%; Score 75; DB 16; Length 129;
Best Local Similarity 100.0%; Pred. No. 7.5e-05;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 LASLSTFQQMWISKQ 15
Db 100 LASLSTFQQMWISKQ 114

RESULT 11
US-10-424-599-143033
; Sequence 143033, Application US/10424599
; Publication No. US20040031072A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J
; APPLICANT: Kovalic, David K
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(53223)B
; CURRENT APPLICATION NUMBER: US/10/424,599
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 285684
; SEQ ID NO 143033
; LENGTH: 133
; TYPE: PRT
; ORGANISM: Glycine max
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT3847_100172C.1.pep
US-10-424-599-143033
```

```
Query Match      100.0%; Score 75; DB 15; Length 133;
Best Local Similarity 100.0%; Pred. No. 7.7e-05;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      1 LASLSTFOQMWSKQ 15
Db      104 LASLSTFOQMWSKQ 118
|||||

RESULT 12
US-10-264-049-4259
; Sequence 4259, Application US/10264049
; Publication No. US20040005579A1
; GENERAL INFORMATION:
; APPLICANT: Birse et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
; FILE REFERENCE: PA133P1
; CURRENT APPLICATION NUMBER: US/10/264,049
; CURRENT FILING DATE: 2002-10-04
; PRIOR APPLICATION NUMBER: PCT/US01/18569
; PRIOR FILING DATE: 2001-06-07
; PRIOR APPLICATION NUMBER: US 60/209,467
; PRIOR FILING DATE: 2000-06-07
; NUMBER OF SEQ ID NOS: 4360
; SOFTWARE: Patent In Ver. 3.1
; SEQ ID NO 4259
; LENGTH: 145
; TYPE: PRT
; ORGANISM: Homo sapiens
; NAME/KEY: MISC_FEATURE
; LOCATION: (3)
; OTHER INFORMATION: Xaa equals any of the twenty naturally occurring L-amino acids
US-10-264-049-4259

Query Match      100.0%; Score 75; DB 15; Length 145;
Best Local Similarity 100.0%; Pred. No. 8.4e-05;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      1 LASLSTFOQMWSKQ 15
Db      116 LASLSTFOQMWSKQ 130
|||||

RESULT 13
US-09-796-692-1239
; Sequence 1239, Application US/09796692
; Publication No. US20020198362A1
; GENERAL INFORMATION:
; APPLICANT: Gaiger, Alexander
; APPLICANT: Mannion, Jane
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE DETECTION, DIAGNOSIS AND THERAPY
; FILE REFERENCE: 2077.001200
; CURRENT APPLICATION NUMBER: US/09/796,692
; CURRENT FILING DATE: 2001-03-01
; PRIOR APPLICATION NUMBER: 60/186,126
; PRIOR FILING DATE: 2000-03-01
; PRIOR APPLICATION NUMBER: 60/190,479
; PRIOR FILING DATE: 2000-03-17
; PRIOR APPLICATION NUMBER: 60/200,999
; PRIOR FILING DATE: 2000-04-27
; PRIOR APPLICATION NUMBER: 60/200,303
; PRIOR FILING DATE: 2000-04-28
; PRIOR APPLICATION NUMBER: 60/200,779
; PRIOR FILING DATE: 2000-04-28
; PRIOR APPLICATION NUMBER: 60/200,999
; PRIOR FILING DATE: 2000-05-01
; PRIOR APPLICATION NUMBER: 60/202,084
; PRIOR FILING DATE: 2000-05-04
; PRIOR APPLICATION NUMBER: 60/206,201
; PRIOR FILING DATE: 2000-05-22
; PRIOR APPLICATION NUMBER: 60/218,950
; PRIOR FILING DATE: 2000-07-14
; PRIOR APPLICATION NUMBER: 60/222,903
; PRIOR FILING DATE: 2000-08-03
; PRIOR APPLICATION NUMBER: 60/223,416
; PRIOR FILING DATE: 2000-08-04
; PRIOR APPLICATION NUMBER: 60/223,378
; PRIOR FILING DATE: 2000-08-07
; NUMBER OF SEQ ID NOS: 9597
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 1696
; LENGTH: 147
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-796-692-1239
```

```
; PRIOR APPLICATION NUMBER: 60/218,950
; PRIOR FILING DATE: 2000-07-14
; PRIOR APPLICATION NUMBER: 60/222,903
; PRIOR FILING DATE: 2000-08-03
; PRIOR APPLICATION NUMBER: 60/223,416
; PRIOR FILING DATE: 2000-08-04
; PRIOR APPLICATION NUMBER: 60/223,378
; PRIOR FILING DATE: 2000-08-07
; NUMBER OF SEQ ID NOS: 9597
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 1239
; LENGTH: 147
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-796-692-1239

Query Match      100.0%; Score 75; DB 9; Length 147;
Best Local Similarity 100.0%; Pred. No. 8.5e-05;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      1 LASLSTFOQMWSKQ 15
Db      118 LASLSTFOQMWSKQ 132
|||||

RESULT 14
US-09-796-692-1696
; Sequence 1696, Application US/09796692
; Publication No. US20020198362A1
; GENERAL INFORMATION:
; APPLICANT: Gaiger, Alexander
; APPLICANT: Mannion, Jane
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE DETECTION, DIAGNOSIS AND THERAPY
; FILE REFERENCE: 2077.001200
; CURRENT APPLICATION NUMBER: US/09/796,692
; CURRENT FILING DATE: 2001-03-01
; PRIOR APPLICATION NUMBER: 60/186,126
; PRIOR FILING DATE: 2000-03-01
; PRIOR APPLICATION NUMBER: 60/190,479
; PRIOR FILING DATE: 2000-03-17
; PRIOR APPLICATION NUMBER: 60/200,545
; PRIOR FILING DATE: 2000-04-27
; PRIOR APPLICATION NUMBER: 60/200,303
; PRIOR FILING DATE: 2000-04-28
; PRIOR APPLICATION NUMBER: 60/200,779
; PRIOR FILING DATE: 2000-04-28
; PRIOR APPLICATION NUMBER: 60/200,999
; PRIOR FILING DATE: 2000-05-01
; PRIOR APPLICATION NUMBER: 60/202,084
; PRIOR FILING DATE: 2000-05-04
; PRIOR APPLICATION NUMBER: 60/206,201
; PRIOR FILING DATE: 2000-05-22
; PRIOR APPLICATION NUMBER: 60/218,950
; PRIOR FILING DATE: 2000-07-14
; PRIOR APPLICATION NUMBER: 60/222,903
; PRIOR FILING DATE: 2000-08-03
; PRIOR APPLICATION NUMBER: 60/223,416
; PRIOR FILING DATE: 2000-08-04
; PRIOR APPLICATION NUMBER: 60/223,378
; PRIOR FILING DATE: 2000-08-07
; NUMBER OF SEQ ID NOS: 9597
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 1696
; LENGTH: 147
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-796-692-1696

Query Match      100.0%; Score 75; DB 9; Length 147;
Best Local Similarity 100.0%; Pred. No. 8.5e-05;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

QY 1 LASLSTFOQMWSKQ 15  
 Db 118 LASLSTFOQMWSKQ 132

RESULT 15  
 US-09-796-692-2250  
 ; Sequence 2250, Application US/09796692  
 ; Publication No. US20020198362A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Gaiger, Alexander  
 ; APPLICANT: Algate, Paul A.  
 ; APPLICANT: Mannion, Jane  
 ; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE DETECTION, DIAGNOSIS AND THERAPY  
 ; TITLE OF INVENTION: HEMATOLOGICAL MALIGNANCIES  
 ; FILE REFERENCE: 2077.001200  
 ; CURRENT APPLICATION NUMBER: US/09/796,692  
 ; CURRENT FILING DATE: 2001-03-01  
 ; PRIOR APPLICATION NUMBER: 60/186,126  
 ; PRIOR FILING DATE: 2000-03-01  
 ; PRIOR APPLICATION NUMBER: 60/190,479  
 ; PRIOR FILING DATE: 2000-03-17  
 ; PRIOR APPLICATION NUMBER: 60/200,545  
 ; PRIOR FILING DATE: 2000-04-27  
 ; PRIOR APPLICATION NUMBER: 60/200,303  
 ; PRIOR FILING DATE: 2000-04-28  
 ; PRIOR APPLICATION NUMBER: 60/200,779  
 ; PRIOR FILING DATE: 2000-04-28  
 ; PRIOR APPLICATION NUMBER: 60/200,999  
 ; PRIOR FILING DATE: 2000-05-01  
 ; PRIOR APPLICATION NUMBER: 60/202,084  
 ; PRIOR FILING DATE: 2000-05-04  
 ; PRIOR APPLICATION NUMBER: 60/206,201  
 ; PRIOR FILING DATE: 2000-05-22  
 ; PRIOR APPLICATION NUMBER: 60/218,950  
 ; PRIOR FILING DATE: 2000-07-14  
 ; PRIOR APPLICATION NUMBER: 60/222,903  
 ; PRIOR FILING DATE: 2000-08-03  
 ; PRIOR APPLICATION NUMBER: 60/223,416  
 ; PRIOR FILING DATE: 2000-08-04  
 ; PRIOR APPLICATION NUMBER: 60/223,378  
 ; PRIOR FILING DATE: 2000-08-07  
 ; NUMBER OF SEQ ID NOS: 9597  
 ; SOFTWARE: FastSeq for Windows Version 3.0  
 ; SEQ ID NO 2250  
 ; LENGTH: 147  
 ; TYPE: PRT  
 ; ORGANISM: Homo sapiens  
 US-09-796-692-2250

Query Match 100.0%; Score 75; DB 9; Length 147;  
 Best Local Similarity 100.0%; Pred. No. 8.5e-05;  
 Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 LASLSTFOQMWSKQ 15  
 Db 118 LASLSTFOQMWSKQ 132

Search completed: April 8, 2005, 12:51:00  
 Job time : 43.6667 secs

Result No.	Query			DB	ID	Description
	Score	Match	Length			
1	80	100.0	71	4	US-09-621-976-6521	Sequence 6521, Ap
2	80	100.0	374	3	US-08-609-236-6	Sequence 6, Appli
3	80	100.0	375	2	US-08-494-151-14	Sequence 14, Appli
4	80	100.0	375	3	US-09-106-217-16	Sequence 16, Appli
5	80	100.0	375	3	US-09-171-337A-7	Sequence 7, Appli
6	80	100.0	375	3	US-09-171-337A-8	Sequence 8, Appli
7	80	100.0	375	4	US-09-631-022-7	Sequence 7, Appli
8	80	100.0	375	4	US-09-631-022-8	Sequence 8, Appli
9	80	100.0	375	4	US-09-976-594-731	Sequence 731, App
10	80	100.0	376	1	US-08-588-113-2	Sequence 2, Appli
11	80	100.0	376	4	US-09-949-016-6100	Sequence 6100, Ap
12	80	100.0	377	3	US-09-106-217-2	Sequence 2, Appli
13	80	100.0	377	4	US-09-919-172-33	Sequence 33, Appli
14	80	100.0	377	4	US-09-917-254-53	Sequence 53, Appli
15	80	100.0	386	4	US-09-949-016-7721	Sequence 7721, Ap
16	80	100.0	399	4	US-09-949-016-9424	Sequence 9424, Ap
17	80	100.0	402	4	US-09-949-016-10757	Sequence 10757, A
18	80	100.0	404	4	US-09-949-016-11313	Sequence 11313, A
19	78	97.5	377	4	US-09-248-796A-14109	Sequence 14109, A
20	74	92.5	41	3	US-09-306-446C-3	Sequence 3, Appli
21	74	92.5	374	3	US-09-306-446C-2	Sequence 2, Appli
22	56	70.0	11	4	US-09-786-066-5	Sequence 5, Appli
23	51	63.7	362	4	US-09-949-016-7725	Sequence 7725, Ap
24	51	63.7	376	4	US-09-538-092-1109	Sequence 1109, Ap
25	51	63.7	376	4	US-09-538-092-1110	Sequence 1110, Ap
26	51	63.7	376	4	US-09-949-016-6656	Sequence 6656, Ap
27	51	63.7	376	4	US-09-949-016-8452	Sequence 8452, Ap



```

Db      26 APRVFPISVGRPRH 40

RESULT 5
US-09-171-337A-7
; Sequence 7, Application US/09171337A
; Patent No. 6300095
; GENERAL INFORMATION:
; APPLICANT: BARREDO FUENTE, Jose Luis
; RODRIGUEZ SAIZ, Marta
; COLLADOS DE LA VIEJA, Alfonso J.
; MORENO VALLE,Miguel Angel
; SALTO MALDONADO, Francisco
; DIEZ GARCIA, Bruno
; TITLE OF INVENTION: PROMOTERS OF THE GENES GLUTAMATE
; DESHYDROGENASE, -N-ACETYLHEXOSAMINIDASE
; AND -ACTIN AND THEIR USE IN FILAMENTOUS
; FUNGI EXPRESSION, SECRETION AND ANTISENSE
; NUMBER OF SEQUENCES: 20
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: LADAS & PARRY
; STREET: 26 WEST 61 STREET
; CITY: NEW YORK
; STATE: NY
; COUNTRY: USA
; ZIP: 10023
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 3-1/4" Disk 1.44MB
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: Microsoft Windows for Workgroups 3.11
; SOFTWARE: WordPerfect 8 for Windows
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/171,337A
; FILING DATE: 14-May-1999
; CLASSIFICATION: 536
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: PCT/ES98/00056
; FILING DATE: 5-MAR-1998
; APPLICATION NUMBER: ES9700482
; FILING DATE: 5-MAR-1997
; ATTORNEY/AGENT INFORMATION:
; NAME: MASS, Clifford J.
; REGISTRATION NUMBER: 30,086
; (C) REF./DOCKET NO.: U-011948-3
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 233288
; INFORMATION FOR SEQ ID NO: 7
; SEQUENCE CHARACTERISTICS:
; LENGTH: 375 amino acids
; TYPE: amino acids
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; ORIGINAL SOURCE:
; ORGANISM: Penicillium chrysogenum
; FEATURE:
; OTHER INFORMATION: amino acid sequence of the -actin
; protein with a molecular weight of
; 41760 Da.
; SEQUENCE DESCRIPTION: SEQ ID NO: 7
US-09-171-337A-7
Query Match      100.0%; Score 80; DB 3; Length 375;
Best Local Similarity 100.0%; Pred. No. 5.7e-06;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 APRVFPISVGRPRH 15
      |||||
Db      26 APRVFPISVGRPRH 40

RESULT 6
US-09-171-337A-8
; Sequence 8, Application US/09171337A
; Patent No. 6300095
; GENERAL INFORMATION:
; APPLICANT: BARREDO FUENTE, Jose Luis
; RODRIGUEZ SAIZ, Marta
; COLLADOS DE LA VIEJA, Alfonso J.
; MORENO VALLE,Miguel Angel
; SALTO MALDONADO, Francisco
; DIEZ GARCIA, Bruno
; TITLE OF INVENTION: PROMOTERS OF THE GENES GLUTAMATE
; DESHYDROGENASE, -N-ACETYLHEXOSAMINIDASE
; AND -ACTIN AND THEIR USE IN FILAMENTOUS
; FUNGI EXPRESSION, SECRETION AND ANTISENSE
; NUMBER OF SEQUENCES: 20
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: LADAS & PARRY
; STREET: 26 WEST 61 STREET
; CITY: NEW YORK
; STATE: NY
; COUNTRY: USA
; ZIP: 10023
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 3-1/4" Disk 1.44MB
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: Microsoft Windows for Workgroups 3.11
; SOFTWARE: WordPerfect 8 for Windows
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/171,337A
; FILING DATE: 14-May-1999
; CLASSIFICATION: 536
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: PCT/ES98/00056
; FILING DATE: 5-MAR-1998
; APPLICATION NUMBER: ES9700482
; FILING DATE: 5-MAR-1997
; ATTORNEY/AGENT INFORMATION:
; NAME: MASS, Clifford J.
; REGISTRATION NUMBER: 30,086
; (C) REF./DOCKET NO.: U-011948-3
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 233288
; INFORMATION FOR SEQ ID NO: 8
; SEQUENCE CHARACTERISTICS:
; LENGTH: 375 amino acids
; TYPE: amino acids
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; ORIGINAL SOURCE:
; ORGANISM: Acremonium chrysogenum
; FEATURE:
; OTHER INFORMATION: amino acid sequence of the -actin
; protein with a molecular weight of
; 41612 Da.
; SEQUENCE DESCRIPTION: SEQ ID NO: 8
US-09-171-337A-8
Query Match      100.0%; Score 80; DB 3; Length 375;
Best Local Similarity 100.0%; Pred. No. 5.7e-06;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 APRVFPISVGRPRH 15
      |||||
Db      26 APRVFPISVGRPRH 40

RESULT 7
US-09-631-022-7
; Sequence 7, Application US/09631022
; Patent No. 6558921
; GENERAL INFORMATION:
; APPLICANT: BARREDO FUENTE, Jose Luis
; RODRIGUEZ SAIZ, Marta

```

COLLADOS DE LA VIEJA, Alfonso J.  
MORENO VALLE Migeul Angel  
SALTO MALDONADO, Francisco  
DIEZ GARCIA, Bruno  
TITLE OF INVENTION: PROMOTERS OF THE GENES GLUTAMATE  
DESHYDROGENASE, -N-ACETYLHEXOSAMINIDASE  
AND -ACTIN AND THEIR USE IN FILAMENTOUS  
FUNGI EXPRESSION, SECRETION AND ANTISENSE  
NUMBER OF SEQUENCES: 20  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: LADAS & PARRY  
STREET: 26 WEST 61 STREET  
CITY: NEW YORK  
STATE: NY  
COUNTRY: USA  
ZIP: 10023  
COMPUTER READABLE FORM:  
MEDIUM TYPE: 3-1/4" Disk 1.44MB  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: Microsoft Windows for Workgroups 3.11  
SOFTWARE: WordPerfect 8 for Windows  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/631,022  
FILING DATE: 02-Aug-2000  
CLASSIFICATION: 536  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 09/171,337  
FILING DATE: 14-MAY-1999  
APPLICATION NUMBER: PCT/ES98/00056  
FILING DATE: 5-MAR-1998  
APPLICATION NUMBER: ES9700482  
FILING DATE: 5-MAR-1997  
ATTORNEY/AGENT INFORMATION:  
NAME: MASS, Clifford J.  
REGISTRATION NUMBER: 30,086  
(C) REF./DOCKET NO.: U-02886-6  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 233288  
INFORMATION FOR SEQ ID NO: 7  
SEQUENCE CHARACTERISTICS:  
LENGTH: 375 amino acids  
TYPE: amino acids  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
ORIGINAL SOURCE:  
ORGANISM: Penicillium chrysogenum  
FEATURE:  
OTHER INFORMATION: amino acid sequence of the -actin  
protein with a molecular weight of  
41760 Da.  
SEQUENCE DESCRIPTION: SEQ ID NO: 7  
US-09-631-022-7  
Query Match 100.0%; Score 80; DB 4; Length 375;  
Best Local Similarity 100.0%; Pred. No. 5.7e-06;  
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 APRAVFPSIVGRPRH 15  
DB 26 APRAVFPSIVGRPRH 40  
RESULT 8  
US-09-631-022-8  
Sequence 8, Application US/09631022  
Patent No. 658921  
GENERAL INFORMATION:  
APPLICANT: BARREDO FUENTE, Jose Luis  
RODRIGUEZ SAIZ, Marta  
COLLADOS DE LA VIEJA, Alfonso J.  
MORENO VALLE, Migeul Angel  
SALTO MALDONADO, Francisco

DIEZ GARCIA, Bruno  
TITLE OF INVENTION: PROMOTERS OF THE GENES GLUTAMATE  
DESHYDROGENASE, -N-ACETYLHEXOSAMINIDASE  
AND -ACTIN AND THEIR USE IN FILAMENTOUS  
FUNGI EXPRESSION, SECRETION AND ANTISENSE  
NUMBER OF SEQUENCES: 20  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: LADAS & PARRY  
STREET: 26 WEST 61 STREET  
CITY: NEW YORK  
STATE: NY  
COUNTRY: USA  
ZIP: 10023  
COMPUTER READABLE FORM:  
MEDIUM TYPE: 3-1/4" Disk 1.44MB  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: Microsoft Windows for Workgroups 3.11  
SOFTWARE: WordPerfect 8 for Windows  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/631,022  
FILING DATE: 02-Aug-2000  
CLASSIFICATION: 536  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 09/171,337  
FILING DATE: 14-MAY-1999  
APPLICATION NUMBER: PCT/ES98/00056  
FILING DATE: 5-MAR-1998  
APPLICATION NUMBER: ES9700482  
FILING DATE: 5-MAR-1997  
ATTORNEY/AGENT INFORMATION:  
NAME: MASS, Clifford J.  
REGISTRATION NUMBER: 30,086  
(C) REF./DOCKET NO.: U-02886-6  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 233288  
INFORMATION FOR SEQ ID NO: 8  
SEQUENCE CHARACTERISTICS:  
LENGTH: 375 amino acids  
TYPE: amino acids  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
ORIGINAL SOURCE:  
ORGANISM: Acremonium chrysogenum  
FEATURE:  
OTHER INFORMATION: amino acid sequence of the -actin  
protein with a molecular weight of  
41612 Da.  
SEQUENCE DESCRIPTION: SEQ ID NO: 8  
US-09-631-022-8  
Query Match 100.0%; Score 80; DB 4; Length 375;  
Best Local Similarity 100.0%; Pred. No. 5.7e-06;  
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 APRAVFPSIVGRPRH 15  
DB 26 APRAVFPSIVGRPRH 40  
RESULT 9  
US-09-976-594-731  
Sequence 731, Application US/09976594  
Patent No. 6673549  
GENERAL INFORMATION:  
APPLICANT: Furness, Michael  
Buchbinder, Jenny  
TITLE OF INVENTION: GENES EXPRESSED IN C3A LIVER CELL CULTURES TREATED WITH STEROIDS  
FILE REFERENCE: PA-0041 US  
CURRENT APPLICATION NUMBER: US/09/976,594  
CURRENT FILING DATE: 2001-10-12  
PRIOR FILING DATE: 2000-10-12



```
; NUMBER OF SEQ ID NOS: 1143
; SOFTWARE: PERL Program
; SEQ ID NO 731
; LENGTH: 375
; TYPE: PRT
; ORGANISM: Homo sapiens
; NAME/KEY: misc feature
; OTHER INFORMATION: Incyte ID No. 6673549 1837317CD1
US-09-976-594-731

Query Match      100.0%; Score 80; DB 4; Length 375;
Best Local Similarity 100.0%; Pred. No. 5.7e-06;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 APRVFPSPVIGRPRH 15
        |||||
Db      26 APRVFPSPVIGRPRH 40
        |||||

RESULT 10
US-08-588-113-2
; Sequence 2, Application US/08588113
; Patent No. 5710003
; GENERAL INFORMATION:
; APPLICANT: McHugh, Kirk M.
; TITLE OF INVENTION: DIAGNOSTIC TEST FOR DETERMINING
; TITLE OF INVENTION: MALIGNANCY OF SMOOTH MUSCLE TUMORS
; NUMBER OF SEQUENCES: 4
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Woodcock Washburn Kurtz Mackiewicz &
; ADDRESSEE: No. 5710003ris
; STREET: One Liberty Place - 46th Floor
; CITY: Philadelphia
; STATE: PA
; COUNTRY: USA
; ZIP: 19103
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/588,113
; FILING DATE:
; CLASSIFICATION: 436
; ATTORNEY/AGENT INFORMATION:
; NAME: Ralph, Rebecca L.
; REGISTRATION NUMBER: 35,152
; REFERENCE/DOCKET NUMBER: TJU-1652
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 215-568-3100
; TELEFAX: 215-568-3439
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 376 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-588-113-2

Query Match      100.0%; Score 80; DB 1; Length 376;
Best Local Similarity 100.0%; Pred. No. 5.7e-06;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 APRVFPSPVIGRPRH 15
        |||||
Db      27 APRVFPSPVIGRPRH 41
        |||||

RESULT 11
US-09-949-016-6100
; Sequence 6100, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; TITLE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CLO01307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 6100
; LENGTH: 376
; TYPE: PRT
; ORGANISM: Human
US-09-949-016-6100

Query Match      100.0%; Score 80; DB 4; Length 376;
Best Local Similarity 100.0%; Pred. No. 5.7e-06;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 APRVFPSPVIGRPRH 15
        |||||
Db      27 APRVFPSPVIGRPRH 41
        |||||

RESULT 12
US-09-106-217-2
; Sequence 2, Application US/09106217
; Patent No. 6063576
; GENERAL INFORMATION:
; APPLICANT: Keating, Mark T.
; APPLICANT: Olson, Timothy M.
; TITLE OF INVENTION: Actin Mutations in Dilated
; TITLE OF INVENTION: Cardiomyopathy, a Heritable Form of Heart Failure
; NUMBER OF SEQUENCES: 18
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Rothwell, Figg, Earnat & Kurz, P.C.
; STREET: 555 Thirteenth Street, N.W., Suite 701 East
; STREET: Tower
; CITY: Washington
; STATE: DC
; COUNTRY: U.S.A.
; ZIP: 20004
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/106,217
; FILING DATE:
; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
; NAME: Saxe, Stephen A.
; REGISTRATION NUMBER: 38,609
; REFERENCE/DOCKET NUMBER: 2323-125
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 202-783-6040
; TELEFAX: 202-783-6031
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 377 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-09-106-217-2
```

Query Match 100.0%; Score 80; DB 3; Length 377;  
Best Local Similarity 100.0%; Pred. No. 5.7e-06;  
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 APRVFPFSIVGRPRH 15  
Db 28 APRVFPFSIVGRPRH 42

## RESULT 13

US-09-919-172-33  
; Sequence 33, Application US/09919172  
; Patent No. 6673545

; GENERAL INFORMATION:

; APPLICANT: Faxis, Mary

; APPLICANT: Turner, Christopher M.

; TITLE OF INVENTION: PROSTATE CANCER MARKERS

; FILE REFERENCE: PA-0036 US

; CURRENT APPLICATION NUMBER: US/09/919,172

; CURRENT FILING DATE: 2001-07-30

; PRIOR APPLICATION NUMBER: 60/222,469

; PRIOR FILING DATE: 2000-07-28

; NUMBER OF SEQ ID NOS: 102

; SOFTWARE: PERL Program

; SEQ ID NO 33

; LENGTH: 377

; TYPE: PRT

; ORGANISM: Homo sapiens

; FEATURE:

; NAME/KEY: misc.feature

; OTHER INFORMATION: Incyte ID No. 6673545 1709118CD1

US-09-919-172-33

Query Match 100.0%; Score 80; DB 4; Length 377;  
Best Local Similarity 100.0%; Pred. No. 5.7e-06;  
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 APRVFPFSIVGRPRH 15  
Db 28 APRVFPFSIVGRPRH 42

## RESULT 14

US-09-917-254-53

; Sequence 53, Application US/09917254

; Patent No. 6703204

; GENERAL INFORMATION:

; APPLICANT: Mutter, George

; APPLICANT: Baak, Jan

; TITLE OF INVENTION: Prognostic Classification of Breast Cancer

; FILE REFERENCE: B0801/7224(JRV)

; CURRENT APPLICATION NUMBER: US/09/917,254

; CURRENT FILING DATE: 2001-07-27

; PRIOR APPLICATION NUMBER: US 60/222,093

; PRIOR FILING DATE: 2000-07-28

; NUMBER OF SEQ ID NOS: 102

; SOFTWARE: PatentIn version 3.0

; SEQ ID NO 53

; LENGTH: 377

; TYPE: PRT

; ORGANISM: Homo Sapiens

US-09-917-254-53

Query Match 100.0%; Score 80; DB 4; Length 377;  
Best Local Similarity 100.0%; Pred. No. 5.7e-06;  
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 APRVFPFSIVGRPRH 15  
Db 28 APRVFPFSIVGRPRH 42

## RESULT 15

US-09-949-016-7721  
; Sequence 7721, Application US/09949016  
; Patent No. 6812339  
; GENERAL INFORMATION:  
; APPLICANT: VENTER, J. Craig et al.  
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED  
; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF  
; FILE REFERENCE: CLO01307  
; CURRENT APPLICATION NUMBER: US/09/949,016  
; CURRENT FILING DATE: 2000-04-14  
; PRIOR APPLICATION NUMBER: 60/241,755  
; PRIOR FILING DATE: 2000-10-20  
; PRIOR APPLICATION NUMBER: 60/237,768  
; PRIOR FILING DATE: 2000-10-03  
; PRIOR APPLICATION NUMBER: 60/231,498  
; PRIOR FILING DATE: 2000-09-08  
; NUMBER OF SEQ ID NOS: 207012  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 7721  
; LENGTH: 386  
; TYPE: PRT  
; ORGANISM: Human  
; US-09-949-016-7721

Query Match 100.0%; Score 80; DB 4; Length 386;  
Best Local Similarity 100.0%; Pred. No. 5.9e-06;  
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 APRVFPFSIVGRPRH 15  
Db 37 APRVFPFSIVGRPRH 51

Search completed: April 8, 2005, 12:07:33  
Job time : 17.5333 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: April 8, 2005, 10:53:18 ; Search time 42.6667 Seconds  
(without alignments)  
116.718 Million cell updates/sec

Title: US-09-423-351C-1

Perfect score: 80

Sequence: 1 APRAVFPISVGRPH 15

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1418010 seqs, 331997259 residues

Total number of hits satisfying chosen parameters: 1418010

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications AA:\*

- 1: /cgn2\_6/ptodata/1/pubpaa/US07\_PUBCOMB.pep:\*
- 2: /cgn2\_6/ptodata/1/pubpaa/PCT\_NEW\_PUB.pep:\*
- 3: /cgn2\_6/ptodata/1/pubpaa/US06\_NEW\_PUB.pep:\*
- 4: /cgn2\_6/ptodata/1/pubpaa/US06\_PUBCOMB.pep:\*
- 5: /cgn2\_6/ptodata/1/pubpaa/US07\_NEW\_PUB.pep:\*
- 6: /cgn2\_6/ptodata/1/pubpaa/PCTUS\_PUBCOMB.pep:\*
- 7: /cgn2\_6/ptodata/1/pubpaa/US08\_NEW\_PUB.pep:\*
- 8: /cgn2\_6/ptodata/1/pubpaa/US08\_PUBCOMB.pep:\*
- 9: /cgn2\_6/ptodata/1/pubpaa/US09\_PUBCOMB.pep:\*
- 10: /cgn2\_6/ptodata/1/pubpaa/US09\_PUBCOMB.pep:\*
- 11: /cgn2\_6/ptodata/1/pubpaa/US09C\_PUBCOMB.pep:\*
- 12: /cgn2\_6/ptodata/1/pubpaa/US09\_NEW\_PUB.pep:\*
- 13: /cgn2\_6/ptodata/1/pubpaa/US10A\_PUBCOMB.pep:\*
- 14: /cgn2\_6/ptodata/1/pubpaa/US10B\_PUBCOMB.pep:\*
- 15: /cgn2\_6/ptodata/1/pubpaa/US10C\_PUBCOMB.pep:\*
- 16: /cgn2\_6/ptodata/1/pubpaa/US10D\_PUBCOMB.pep:\*
- 17: /cgn2\_6/ptodata/1/pubpaa/US10\_NEW\_PUB.pep:\*
- 18: /cgn2\_6/ptodata/1/pubpaa/US11\_NEW\_PUB.pep:\*
- 19: /cgn2\_6/ptodata/1/pubpaa/US60\_NEW\_PUB.pep:\*
- 20: /cgn2\_6/ptodata/1/pubpaa/US60\_PUBCOMB.pep:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	80	100.0	105	15	US-10-264-049-3601
2	80	100.0	135	16	US-10-767-701-47318
3	80	100.0	136	15	US-10-424-599-280811
4	80	100.0	142	15	US-10-424-599-184471
5	80	100.0	197	15	US-10-424-599-280810
6	80	100.0	201	15	US-10-264-049-4308
7	80	100.0	204	15	US-10-264-049-4148
8	80	100.0	208	15	US-10-424-599-282905
9	80	100.0	219	14	US-10-002-631C-56
10	80	100.0	230	16	US-10-767-701-42770
11	80	100.0	342	15	US-10-108-260A-4008
12	80	100.0	347	15	US-10-424-599-283333
13	80	100.0	371	16	US-10-322-281-68

Sequence 93, Appl	14	US-10-205-194-93	375	14	Sequence 52, Appl
Sequence 88, Appl	15	US-10-316-253-88	375	14	Sequence 64, Appl
Sequence 1786, Ap	16	US-10-369-493-1786	375	15	Sequence 162009,
Sequence 2436, Ap	17	US-10-369-493-2436	375	15	Sequence 222816,
Sequence 5927, Ap	18	US-10-369-493-5927	375	15	Sequence 223492,
Sequence 94, Appl	19	US-10-205-331-94	375	15	Sequence 283336,
Sequence 82, Appl	20	US-10-260-708-82	375	15	Sequence 66124, A
Sequence 1, Appl	21	US-10-477-369-1	375	16	Sequence 4, Appl
Sequence 101, App	22	US-10-341-434-101	376	15	Sequence 121952,
Sequence 179917,	23	US-10-437-963-179917	376	16	Sequence 148877,
Sequence 202420,	24	US-10-437-963-202420	376	16	Sequence 198295,
Sequence 45677, A	25	US-10-767-701-45677	376	16	Sequence 47239, A
Sequence 46090, A	26	US-10-767-701-46090	376	16	Sequence 52458, A
Sequence 33, Appl	27	US-09-919-172-33	377	9	Sequence 59265, A
Sequence 52, Appl	28	US-10-338-777-52	377	14	Sequence 59967, A
Sequence 64, Appl	29	US-10-236-031B-64	377	15	Sequence 61003, A
Sequence 162009,	30	US-10-424-599-162009	377	15	Sequence 62783, A
Sequence 222816,	31	US-10-424-599-222816	377	15	
Sequence 223492,	32	US-10-424-599-223492	377	15	
Sequence 283336,	33	US-10-424-599-283336	377	15	
Sequence 66124, A	34	US-10-425-114-66124	377	15	
Sequence 4, Appl	35	US-10-333-680-4	377	16	
Sequence 121952,	36	US-10-437-963-121952	377	16	
Sequence 148877,	37	US-10-437-963-148877	377	16	
Sequence 198295,	38	US-10-437-963-198295	377	16	
Sequence 47239, A	39	US-10-767-701-47239	377	16	
Sequence 52458, A	40	US-10-425-114-52458	378	15	
Sequence 59265, A	41	US-10-425-114-59265	378	15	
Sequence 59967, A	42	US-10-425-114-59967	378	15	
Sequence 61003, A	43	US-10-425-114-61003	378	15	
Sequence 62783, A	44	US-10-425-114-62783	378	15	
	45	US-10-425-114-62783	378	15	

## ALIGNMENTS

RESULT 1  
US-10-264-049-3601  
; Sequence 3601, Application US/10264049  
; Publication No. US20040005579A1  
; GENERAL INFORMATION:  
; APPLICANT: Birse et al.  
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies  
; FILE REFERENCE: PA133PI  
; CURRENT APPLICATION NUMBER: US/10/264,049  
; CURRENT FILING DATE: 2002-10-04  
; PRIOR APPLICATION NUMBER: PCT/US01/18569  
; PRIOR FILING DATE: 2001-06-07  
; PRIOR APPLICATION NUMBER: US 60/209,467  
; PRIOR FILING DATE: 2000-06-07  
; NUMBER OF SEQ ID NOS: 4360  
; SOFTWARE: Patent In Ver. 3.1  
; SEQ ID NO 3601  
; LENGTH: 105  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
; FEATURE:  
; NAME/KEY: MISC FEATURE  
; LOCATION: (103)  
; OTHER INFORMATION: Xaa equals any of the twenty naturally occurring L-amino acids  
US-10-264-049-3601

Query Match 100.0%; Score 80; DB 15; Length 105;  
Best Local Similarity 100.0%; Pred. No. 1.7e-05;  
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 APRAVFPISVGRPH 15  
DB 27 APRAVFPISVGRPH 41

RESULT 2

US-10-767-701-47318  
; Sequence 47318, Application US/10767701  
; Publication No. US20040172684A1  
; GENERAL INFORMATION:  
; APPLICANT: Kovalic, David K.  
; APPLICANT: Zhou, Yihua  
; APPLICANT: Cao, Yongwei  
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With  
; TITLE OF INVENTION: Plants and Uses Thereof For Plant Improvement  
; FILE REFERENCE: 38-21(53223)B  
; CURRENT APPLICATION NUMBER: US/10/767,701  
; CURRENT FILING DATE: 2004-01-29  
; NUMBER OF SEQ ID NOS: 63128  
; SEQ ID NO 47318  
; LENGTH: 135  
; TYPE: PRT  
; ORGANISM: Sorghum bicolor  
; FEATURE:  
; NAME/KEY: unsure  
; LOCATION: (1)..(135)  
; OTHER INFORMATION: unsure at all Xaa locations  
; OTHER INFORMATION: Clone ID: SORBI-28MAY03-C105\_2.pep  
US-10-767-701-47318

Query Match 100.0%; Score 80; DB 16; Length 135;  
Best Local Similarity 100.0%; Pred. No. 2.1e-05;  
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 APRAVPSIVGRPRH 15  
Db 28 APRAVPSIVGRPRH 42

RESULT 3  
US-10-424-599-280811  
; Sequence 280811, Application US/10424599  
; Publication No. US20040031072A1  
; GENERAL INFORMATION:  
; APPLICANT: La Rosa Thomas J  
; APPLICANT: Kovalic David K  
; APPLICANT: Zhou Yihua  
; APPLICANT: Cao Yongwei  
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With  
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement  
; FILE REFERENCE: 38-21(53223)B  
; CURRENT APPLICATION NUMBER: US/10/424,599  
; CURRENT FILING DATE: 2003-04-28  
; NUMBER OF SEQ ID NOS: 285684  
; SEQ ID NO 280811  
; LENGTH: 136  
; TYPE: PRT  
; ORGANISM: Glycine max  
; FEATURE:  
; OTHER INFORMATION: Clone ID: PAT\_MRT3847\_95595C.1.pep  
US-10-424-599-280811

Query Match 100.0%; Score 80; DB 15; Length 136;  
Best Local Similarity 100.0%; Pred. No. 2.2e-05;  
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 APRAVPSIVGRPRH 15  
Db 28 APRAVPSIVGRPRH 42

RESULT 4  
US-10-424-599-184471  
; Sequence 184471, Application US/10424599  
; Publication No. US20040031072A1  
; GENERAL INFORMATION:  
; APPLICANT: La Rosa Thomas J  
; APPLICANT: Kovalic David K

; APPLICANT: Zhou Yihua  
; APPLICANT: Cao Yongwei  
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With  
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement  
; FILE REFERENCE: 38-21(53223)B  
; CURRENT APPLICATION NUMBER: US/10/424,599  
; CURRENT FILING DATE: 2003-04-28  
; NUMBER OF SEQ ID NOS: 285684  
; SEQ ID NO 184471  
; LENGTH: 142  
; TYPE: PRT  
; ORGANISM: Glycine max  
; FEATURE:  
; OTHER INFORMATION: Clone ID: PAT\_MRT3847\_137594C.1.pep  
US-10-424-599-184471

Query Match 100.0%; Score 80; DB 15; Length 142;  
Best Local Similarity 100.0%; Pred. No. 2.3e-05;  
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 APRAVPSIVGRPRH 15  
Db 35 APRAVPSIVGRPRH 49

RESULT 5  
US-10-424-599-280810  
; Sequence 280810, Application US/10424599  
; Publication No. US20040031072A1  
; GENERAL INFORMATION:  
; APPLICANT: La Rosa Thomas J  
; APPLICANT: Kovalic David K  
; APPLICANT: Zhou Yihua  
; APPLICANT: Cao Yongwei  
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With  
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement  
; FILE REFERENCE: 38-21(53223)B  
; CURRENT APPLICATION NUMBER: US/10/424,599  
; CURRENT FILING DATE: 2003-04-28  
; NUMBER OF SEQ ID NOS: 285684  
; SEQ ID NO 280810  
; LENGTH: 197  
; TYPE: PRT  
; ORGANISM: Glycine max  
; FEATURE:  
; NAME/KEY: unsure  
; LOCATION: (1)..(197)  
; OTHER INFORMATION: unsure at all Xaa locations  
; OTHER INFORMATION: Clone ID: PAT\_MRT3847\_95594C.1.pep  
US-10-424-599-280810

Query Match 100.0%; Score 80; DB 15; Length 197;  
Best Local Similarity 100.0%; Pred. No. 3.2e-05;  
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 APRAVPSIVGRPRH 15  
Db 28 APRAVPSIVGRPRH 42

RESULT 6  
US-10-264-049-4308  
; Sequence 4308, Application US/10264049  
; Publication No. US20040005579A1  
; GENERAL INFORMATION:  
; APPLICANT: Birse et al.  
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies  
; FILE REFERENCE: PAL33P1  
; CURRENT APPLICATION NUMBER: US/10/264,049  
; CURRENT FILING DATE: 2002-10-04  
; PRIOR APPLICATION NUMBER: PCT/US01/18569  
; PRIOR FILING DATE: 2001-06-07

```

US-10-002-631C-56
; Sequence 56, Application US/10002631C
; Publication No. US20030157486A1
; GENERAL INFORMATION:
; APPLICANT: Graff, Jonathon M.
; APPLICANT: Muenster, Matthew
; TITLE OF INVENTION: METHODS TO IDENTIFY SIGNAL SEQUENCES
; FILE REFERENCE: A34943 090495.0243
; CURRENT APPLICATION NUMBER: US/10/002,631C
; CURRENT FILING DATE: 2001-10-31
; PRIOR APPLICATION NUMBER: 60/300,309
; PRIOR FILING DATE: 2001-06-21
; NUMBER OF SEQ ID NOS: 324
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 56

```

```
; LENGTH: 219
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-002-631C-56

Query Match      100.0%; Score 80; DB 14; Length 219;
Best Local Similarity 100.0%; Pred. No. 3.6e-05;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 APRAVFPSIVGRPRH 15
Db 53 APRAVFPSIVGRPRH 67
|||||

RESULT 10
US-10-767-701-42770
; Sequence 42770, Application US/10767701
; Publication No. US20040172684A1
; GENERAL INFORMATION:
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(53535)B
; CURRENT APPLICATION NUMBER: US/10/767,701
; CURRENT FILING DATE: 2004-01-29
; NUMBER OF SEQ ID NOS: 63128
; SEQ ID NO 42770
; LENGTH: 230
; TYPE: PRT
; ORGANISM: Sorghum bicolor
; FEATURE:
; OTHER INFORMATION: Clone ID: SORBI-28MAY03-C17410_1.pep
US-10-767-701-42770

Query Match      100.0%; Score 80; DB 16; Length 230;
Best Local Similarity 100.0%; Pred. No. 3.8e-05;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 APRAVFPSIVGRPRH 15
Db 28 APRAVFPSIVGRPRH 42
|||||

RESULT 11
US-10-108-260A-4008
; Sequence 4008, Application US/10108260A
; Publication No. US20040005560A1
; GENERAL INFORMATION:
; APPLICANT: HELIX RESEARCH INSTITUTE
; TITLE OF INVENTION: No. US20040005560A1el full length cDNA
; FILE REFERENCE: H1-A0106
; CURRENT APPLICATION NUMBER: US/10/108,260A
; CURRENT FILING DATE: 2002-03-27
; NUMBER OF SEQ ID NOS: 5458
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 4008
; LENGTH: 342
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-108-260A-4008

Query Match      100.0%; Score 80; DB 15; Length 342;
Best Local Similarity 100.0%; Pred. No. 5.7e-05;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 APRAVFPSIVGRPRH 15
Db 28 APRAVFPSIVGRPRH 42
|||||

RESULT 12
US-10-424-599-283333
; Sequence 283333, Application US/10424599
; Publication No. US20040031072A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa Thomas J
; APPLICANT: Kovalic David K
; APPLICANT: Zhou Yihua
; APPLICANT: Cao Yongwei
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(53223)B
; CURRENT APPLICATION NUMBER: US/10/424,599
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 285684
; SEQ ID NO 283333
; LENGTH: 347
; TYPE: PRT
; ORGANISM: Glycine max
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (1)..(347)
; OTHER INFORMATION: unsure at all Xaa locations
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT3847_97873C.1.pep
US-10-424-599-283333

Query Match      100.0%; Score 80; DB 15; Length 347;
Best Local Similarity 100.0%; Pred. No. 5.8e-05;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 APRAVFPSIVGRPRH 15
Db 28 APRAVFPSIVGRPRH 42
|||||

RESULT 13
US-10-322-281-68
; Sequence 68, Application US/10322281
; Publication No. US20040126762A1
; GENERAL INFORMATION:
; APPLICANT: David W. Morris
; APPLICANT: Marc S. Malandro
; TITLE OF INVENTION: Novel Compositions and Methods in Cancer
; FILE REFERENCE: 529452001000
; CURRENT APPLICATION NUMBER: US/10/322,281
; CURRENT FILING DATE: 2002-12-17
; NUMBER OF SEQ ID NOS: 866
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 68
; LENGTH: 371
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-322-281-68

Query Match      100.0%; Score 80; DB 16; Length 371;
Best Local Similarity 100.0%; Pred. No. 6.2e-05;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 APRAVFPSIVGRPRH 15
Db 26 APRAVFPSIVGRPRH 40
|||||

RESULT 14
US-10-205-194-93
; Sequence 93, Application US/10205194
; Publication No. US20030134301A1
; GENERAL INFORMATION:
; APPLICANT: Warner-Lambert Company
; APPLICANT: Lee, Kevin
; APPLICANT: Dixon, Alistair
; APPLICANT: Brooksbank, Robert
; APPLICANT: Pinnoch, Robert
```

; TITLE OF INVENTION: Identification and Use of Molecules Implicated in Pain  
 ; FILE REFERENCE: WL-A-018201  
 ; CURRENT APPLICATION NUMBER: US/10/205,194  
 ; CURRENT FILING DATE: 5200-07-24  
 ; PRIOR APPLICATION NUMBER: GB 0118354.0  
 ; PRIOR FILING DATE: 2001-07-27  
 ; NUMBER OF SEQ ID NOS: 177  
 ; SOFTWARE: PatentIn Ver. 2.1  
 ; SEQ ID NO 93  
 ; LENGTH: 375  
 ; TYPE: PRT  
 ; ORGANISM: Rattus rattus  
 ; FEATURE:  
 ; OTHER INFORMATION: Cytoplasmic gamma isoform of actin  
 US-10-205-194-93

Query Match 100.0%; Score 80; DB 14; Length 375;  
 Best Local Similarity 100.0%; Pred. No. 6.3e-05;  
 Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 APRAVFPSIVGRPRH 15  
 |||||  
 DB 26 APRAVFPSIVGRPRH 40

RESULT 15  
 US-10-316-253-88  
 ; Sequence 88, Application US/10316253  
 ; Publication No. US20030162706A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: The Procter & Gamble Company  
 ; APPLICANT: Peters, Kevin  
 ; APPLICANT: Thompson, Larry  
 ; APPLICANT: Wang, Feng  
 ; APPLICANT: Greis, Kenneth  
 ; TITLE OF INVENTION: Angiogenesis Modulating Proteins  
 ; FILE REFERENCE: 8865M  
 ; CURRENT APPLICATION NUMBER: US/10/316,253  
 ; CURRENT FILING DATE: 2002-12-10  
 ; PRIOR APPLICATION NUMBER: US 60/355,295  
 ; PRIOR FILING DATE: 2002-02-08  
 ; NUMBER OF SEQ ID NOS: 308  
 ; SOFTWARE: PatentIn version 3.1  
 ; SEQ ID NO 88  
 ; LENGTH: 375  
 ; TYPE: PRT  
 ; ORGANISM: Rattus norvegicus  
 US-10-316-253-88

Query Match 100.0%; Score 80; DB 14; Length 375;  
 Best Local Similarity 100.0%; Pred. No. 6.3e-05;  
 Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 APRAVFPSIVGRPRH 15  
 |||||  
 DB 26 APRAVFPSIVGRPRH 40

Search completed: April 8, 2005, 12:50:57  
 Job time : 43.6667 secs

**THIS PAGE BLANK (USPTO)**



GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: April 8, 2005, 09:58:56 ; Search time 9.33333 Seconds  
(without alignments)  
154.634 Million cell updates/sec

Title: US-09-423-351C-1

Perfect score: 80

Sequence: 1 APRAVFPISIVGRPRH 15

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 283416 seqs, 96216763 residues

Total number of hits satisfying chosen parameters: 283416

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

PIR\_79:\*

1: pir1:\*

2: pir2:\*

3: pir3:\*

4: pir4:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	80	100.0	43	2 C43616	actin alpha, cardi
2	80	100.0	86	2 C43616	actin beta, cyto
3	80	100.0	137	2 A28258	actin 5C - fruit f
4	80	100.0	140	2 S03109	actin - pin mould
5	80	100.0	158	2 I49465	alpha-cardiac acti
6	80	100.0	195	2 S20097	actin 85C - potato
7	80	100.0	213	2 A61043	actin CA15 - sea s
8	80	100.0	308	2 A03000	actin 3 - fruit fl
9	80	100.0	328	2 S05430	actin beta - grass
10	80	100.0	336	2 T04085	actin - maize (fira
11	80	100.0	362	2 A26559	actin type 5, cyto
12	80	100.0	370	2 A29664	actin - sea urchin
13	80	100.0	374	1 ATE08	actin beta - bovin
14	80	100.0	374	1 ATE08	actin gamma - bovi
15	80	100.0	374	2 JC5818	gamma-actin - huma
16	80	100.0	375	1 ATEBOM	actin, aortic smoo
17	80	100.0	375	1 ATRB	actin, skeletal mu
18	80	100.0	375	1 ATRTC	actin beta - rat
19	80	100.0	375	1 A48324	actin beta, cytosk
20	80	100.0	375	1 ATRAX	actin - Acanthamo
21	80	100.0	375	1 ATRB	actin - yeast (Sac
22	80	100.0	375	1 ATRCH	actin beta - chick
23	80	100.0	375	1 ATRDO	actin - slime mold
24	80	100.0	375	1 ATRHUB	actin beta - human
25	80	100.0	375	1 ATRHUG	actin gamma 1 - hu
26	80	100.0	375	1 ATRMSB	actin gamma - mous
27	80	100.0	375	1 ATRMSG	actin gamma - mous
28	80	100.0	375	1 ATRBB	actin beta, non-mu
29	80	100.0	375	1 ATRM1	actin - maize

30 80 100.0 375 1 JS0702 actin - yeast (Sac  
31 80 100.0 375 1 S11222 actin gamma, cytos  
32 80 100.0 375 2 S33386 actin, cytosolic (  
33 80 100.0 375 2 JT0385 actin gamma - Emer  
34 80 100.0 375 2 S47897 actin 1 - Pneumocy  
35 80 100.0 375 2 T25272 hypothetical prote  
36 80 100.0 375 2 A32798 actin - yeast (Klu  
37 80 100.0 375 2 A26836 actin - fission ye  
38 80 100.0 375 2 S71125 actin beta-2, cyto  
39 80 100.0 375 2 S71124 actin beta-1, cyto  
40 80 100.0 375 2 S71126 actin beta, cyto  
41 80 100.0 375 2 A55001 actin beta - goose  
42 80 100.0 375 2 A54728 actin alpha, cardi  
43 80 100.0 375 2 S03126 actin - imperfect  
44 80 100.0 375 2 S70377 actin - Phaffia rh  
45 80 100.0 376 1 ATFF7 actin 7 - fruit fl

#### ALIGNMENTS

##### RESULT 1

A43616  
actin alpha, cardiac - chicken (fragment)  
C:Species: Gallus gallus (chicken)  
C:Date: 11-Dec-1992 #sequence\_revision 11-Dec-1992 #text\_change 16-Jul-1999  
C:Accession: A43616  
R:Paterson, B.M.; Eldridge, J.D.  
Science 224, 1436-1438, 1984  
A:Title: alpha-Cardiac actin is the major sarcomeric isoform expressed in embryonic avia  
A:Reference number: A43616; MUID:84223949; PMID:6729461  
A:Accession: A43616  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-43 <PAT>  
A:Cross-references: GB:M10607  
C:Superfamily: actin  
C:Keywords: cardiac muscle; heart

Query Match 100.0%; Score 80; DB 2; Length 43;  
Best Local Similarity 100.0%; Pred. No. 6.2e-07;  
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 APRAVFPISIVGRPRH 15  
Db 28 APRAVFPISIVGRPRH 42  
|||||

##### RESULT 2

A43616  
actin beta, cytosolic - chicken (fragment)  
C:Species: Gallus gallus (chicken)  
C:Date: 11-Dec-1992 #sequence\_revision 11-Dec-1992 #text\_change 09-Jul-2004  
C:Accession: C43616  
R:Paterson, B.M.; Eldridge, J.D.  
Science 224, 1436-1438, 1984  
A:Title: alpha-Cardiac actin is the major sarcomeric isoform expressed in embryonic avia  
A:Reference number: A43616; MUID:84223949; PMID:6729461  
A:Accession: C43616  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-86 <PAT>  
A:Cross-references: UNIPROT:Q90736; GB:K02173; NID:g211054; PIDN:AAA98513.1; PID:g211055  
C:Superfamily: actin  
C:Keywords: cytosol; methylated amino acid  
P:73/Modified site: 3'-methylhistidine (His) #status predicted

Query Match 100.0%; Score 80; DB 2; Length 86;  
Best Local Similarity 100.0%; Pred. No. 1.3e-06;  
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 APRAVFPISIVGRPRH 15  
|||||

Db 26 APRAVFPSIVGRPRH 40

RESULT 3

A28258

actin 5C - fruit fly (*Drosophila melanogaster*)

C:Species: *Drosophila melanogaster*

C>Date: 30-Sep-1989 #sequence\_revision 30-Sep-1989 #text\_change 09-Jul-2004

C:Accession: A28258

R:Vigoreaux, J.O.; Tobin, S.L.

Genes Dev. 1, 1161-1171, 1987

A:Title: Stage-specific selection of alternative transcriptional initiation sites from

A:Reference number: A28258; MUID:88112795; PMID:3123314

A:Accession: A28258

A:Molecule type: mRNA

A:Residues: 1-137 <VIG>

A:CROSS-references: UNIPROT:P10987

A:Note: the authors translated the codon GAG for residue 96 as Gly

C:Genetics:

A:Gene: FlyBase:Act5C

A:CROSS-references: FlyBase:FBgn0000042

C:Superfamily: actin

C:Keywords: methylated amino acid

F,74/Modified site: 3'-methylhistidine (His) #status predicted

Query Match 100.0%; Score 80; DB 2; Length 137;

Best Local Similarity 100.0%; Pred. No. 2.1e-06;

Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 APRAVFPSIVGRPRH 15

Db 27 APRAVFPSIVGRPRH 41

RESULT 4

S03109

actin - pin mould (*Absidia glauca*) (fragment)

C:Species: *Absidia glauca*

C>Date: 07-Sep-1990 #sequence\_revision 07-Sep-1990 #text\_change 09-Jul-2004

C:Accession: S03109

R:Burmester, A.; Weigel, C.; Woesteneyer, J.

submitted to the EMBL Data Library, June 1988

A:Reference number: S03109

A:Accession: S03109

A:Molecule type: DNA

A:Residues: 1-140 <BUR>

A:CROSS-references: UNIPROT:P10982; EMBL:X07999; NID:G2303; PIDN:CAA30804.1; PID:g578097

C:Genetics:

A:Gene: act1

A:Introns: 9/3; 31/3

C:Superfamily: actin

C:Keywords: methylated amino acid

F,75/Modified site: 3'-methylhistidine (His) #status predicted

Query Match 100.0%; Score 80; DB 2; Length 140;

Best Local Similarity 100.0%; Pred. No. 2.1e-06;

Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 APRAVFPSIVGRPRH 15

Db 28 APRAVFPSIVGRPRH 42

RESULT 5

I49465

alpha-cardiac actin - mouse (fragment)

C:Species: *Mus musculus* (house mouse)

C>Date: 02-Jul-1996 #sequence\_revision 02-Jul-1996 #text\_change 09-Jul-2004

C:Accession: I49465; I49466

R:Garner, I.; Minty, A.J.; Alonso, S.; Barton, P.J.R.; Buckingham, M.E.

EMBO J. 5, 2559-2567, 1986

A:Title: A 5' duplication of the alpha-cardiac actin gene in BALB/c mice is associated

A:Reference number: I49465; MUID:87053822; PMID:3023046

A:Accession: I49465

A:Status: translated from GB/EMBL/DDBJ

A:Molecule type: DNA

A:Residues: 1-158 <RES>

A:CROSS-references: UNIPROT:Q61273; GB:M26775; NID:G191646; PIDN:AAA37165.1; PID:g553858

A:Accession: I49466

A:Status: translated from GB/EMBL/DDBJ

A:Molecule type: DNA

A:Residues: 8-158 <RE2>

A:CROSS-references: GB:M26776; NID:G191649; PIDN:AAA37166.1; PID:g553859

A:Experimental source: adult cardiac muscle, BALB/c mice

C:Genetics:

A:Introns: 50/3

C:Superfamily: actin

C:Keywords: cardiac muscle; heart

Query Match 100.0%; Score 80; DB 2; Length 158;

Best Local Similarity 100.0%; Pred. No. 2.4e-06;

Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 APRAVFPSIVGRPRH 15

Db 35 APRAVFPSIVGRPRH 49

RESULT 6

S20097

actin 85c - potato (fragment)

C:Species: *Solanum tuberosum* (potato)

C>Date: 22-Nov-1993 #sequence\_revision 17-Oct-1997 #text\_change 09-Jul-2004

C:Accession: S20097

R:Drouin, G.; Dover, G.A.

J. Mol. Evol. 31, 132-150, 1990

A:Title: Independent gene evolution in the potato actin gene family demonstrated by phyl

A:Reference number: S20092; MUID:91012599; PMID:2120451

A:Accession: S20097

A:Molecule type: DNA

A:Residues: 1-195 <DRO>

A:CROSS-references: UNIPROT:P30170; EMBL:X55747; NID:G21541; PIDN:CAA39277.1; PID:G13455

C:Genetics:

A:Introns: 132/1

C:Superfamily: actin

C:Keywords: cytoskeleton; structural protein

Query Match 100.0%; Score 80; DB 2; Length 195;

Best Local Similarity 100.0%; Pred. No. 3e-06;

Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 APRAVFPSIVGRPRH 15

Db 8 APRAVFPSIVGRPRH 22

RESULT 7

A61043

actin CA15 - sea squirt (*Styela clava*) (fragments)

C:Species: *Styela clava*

C>Date: 31-Dec-1993 #sequence\_revision 31-Dec-1993 #text\_change 09-Jul-2004

C:Accession: A61043

R:Beach, R.L.; Jeffery, W.R.

Dev. Genet. 11, 2-14, 1990

A:Title: Temporal and spatial expression of a cytoskeletal actin gene in the ascidian St

A:Reference number: A61043; MUID:90298580; PMID:2361333

A:Accession: A61043

A:Status: not compared with conceptual translation

A:Molecule type: mRNA

A:Residues: 1-213 <BEA>

A:CROSS-references: UNIPROT:Q7M3Y7

C:Comment: This sequence is expressed in cells undergoing rapid cell division.

C:Superfamily: actin

C:Keywords: cytoskeleton; methylated amino acid; mitosis; structural protein

F,73/Modified site: 3'-methylhistidine (His) #status predicted

```
Query Match      100.0%; Score 80; DB 2; Length 213;
Best Local Similarity 100.0%; Pred. No. 3.3e-06;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 APRAVFPSIVGRPRH 15
DB      26 APRAVFPSIVGRPRH 40
      |||||
      |||||
      |||||

RESULT 8
A03000
actin 3 - fruit fly (Drosophila melanogaster) (fragments)
C:Species: Drosophila melanogaster
C>Date: 31-Mar-1991 #sequence_revision 31-Mar-1991 #text_change 09-Jul-2004
C:Accession: A03000
R:Fyrborg, E.A.; Bond, B.J.; Hershey, N.D.; Mixter, K.S.; Davidson, N.
Call 24, 107-116, 1981
A>Title: The actin genes of Drosophila: protein coding regions are highly conserved but
A:Reference number: A03000; MUID:81210174; PMID:6263481
A:Accession: A03000
A:Molecule type: DNA
A:Residues: 1-308 <FVR>
A:Cross-references: UNIPROT:P02572
A>Note: there are 68 unsequenced residues between positions 160 and 161. Partial sequenc
C:Genetics:
A:Gene: FlyBase:Act42A
A:Cross-references: FlyBase:FBgn0000043
A:Map position: 42A
C:Superfamily: actin
C:Keywords: methylated amino acid
F,7/4/Modified site: 3'-methylhistidine (His) #status predicted

Query Match      100.0%; Score 80; DB 2; Length 308;
Best Local Similarity 100.0%; Pred. No. 4.8e-06;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 APRAVFPSIVGRPRH 15
DB      27 APRAVFPSIVGRPRH 41
      |||||
      |||||
      |||||

RESULT 9
S05430
actin beta - grass carp
C:Species: Ctenopharyngodon idella (grass carp)
C>Date: 07-Jun-1990 #sequence_revision 07-Jun-1990 #text_change 09-Jul-2004
C:Accession: S05430
R:Liu, Z.; Zhu, Z.; Roberg, K.; Faras, A.J.; Guise, K.S.; Kapuscinski, A.R.; Hackett, P.
Nucleic Acids Res. 17, 5850, 1989
A>Title: The beta-actin gene of carp (Ctenopharyngodon idella).
A:Reference number: S05430; MUID:89345185; PMID:2762162
A:Accession: S05430
A>Status: translation not shown
A:Molecule type: DNA
A:Residues: 1-328 <LIU>
A:Cross-references: UNIPROT:P83751; EMBL:M25013
C:Genetics:
A:introns: 41/3; 121/3; 268/1
C:Superfamily: actin
C:Keywords: cytoskeleton; methylated amino acid
F,7/3/Modified site: 3'-methylhistidine (His) #status predicted

Query Match      100.0%; Score 80; DB 2; Length 328;
Best Local Similarity 100.0%; Pred. No. 5.2e-06;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 APRAVFPSIVGRPRH 15
DB      26 APRAVFPSIVGRPRH 40
      |||||
      |||||
      |||||

RESULT 10
```

```
T04085
actin - maize (fragment)
C:Species: Zea mays (maize)
C>Date: 23-Apr-1999 #sequence_revision 23-Apr-1999 #text_change 09-Jul-2004
C:Accession: T04085
R:de Sa, M.; Drouin, G.
Mol. Biol. Evol. 13, 1198-1212, 1996
A>Title: Phylogeny and substitution rates of angiosperm actin genes.
A:Reference number: Z15197; MUID:97051711; PMID:8896372
A:Accession: T04085
A>Status: preliminary; translated from GB/EMBL/DDBBJ
A:Molecule type: DNA
A:Residues: 1-336 <DES>
A:Cross-references: UNIPROT:P93638; EMBL:U60513; NID:g1498392; PIDN:AAB40107.1; PID:g1449
C:Genetics:
A:Gene: Maz63
A>Note: Intron positions not resolved (incomplete sequence)
C:Superfamily: actin
C:Keywords: cytoskeleton; structural protein

Query Match      100.0%; Score 80; DB 2; Length 336;
Best Local Similarity 100.0%; Pred. No. 5.3e-06;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 APRAVFPSIVGRPRH 15
DB      8 APRAVFPSIVGRPRH 22
      |||||
      |||||
      |||||

RESULT 11
A26559
actin type 5, cytosolic - chicken
C:Species: Gallus gallus (chicken)
C>Date: 09-Sep-1987 #sequence_revision 09-Sep-1987 #text_change 05-Dec-1997
C:Accession: A26559
R:Bergsma, D.J.; Chang, K.S.; Schwartz, R.J.
Mol. Cell. Biol. 5, 1151-1162, 1985
A:Reference number: A26559; MUID:85213487; PMID:4000121
A:Accession: A26559
A:Molecule type: DNA
A:Residues: 1-362 <BER>
C:Superfamily: actin
C:Keywords: cytosol; methylated amino acid
F,7/4/Modified site: 3'-methylhistidine (His) #status predicted

Query Match      100.0%; Score 80; DB 2; Length 362;
Best Local Similarity 100.0%; Pred. No. 5.7e-06;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 APRAVFPSIVGRPRH 15
DB      27 APRAVFPSIVGRPRH 41
      |||||
      |||||
      |||||

RESULT 12
A29664
actin - sea urchin (Strongylocentrotus purpuratus)
C:Species: Strongylocentrotus purpuratus (purple urchin)
C>Date: 15-Dec-1988 #sequence_revision 30-Sep-1991 #text_change 05-Dec-1997
C:Accession: A29664
R:Crain Jr., W.R.; Boshart, M.F.; Cooper, A.D.; Durica, D.S.; Nagy, A.; Steffen, D.
J. Mol. Evol. 25, 37-45, 1987
A>Title: The sequence of a sea urchin muscle actin gene suggests a gene conversion with
A:Reference number: A29664; MUID:87311761; PMID:3114500
A:Accession: A29664
A:Molecule type: DNA
A:Residues: 1-370 <CRA>
A:Cross-references: GB:X05739; GB:X05740; GB:X05741; GB:X05742; GB:X05743
A>Note: the authors translated the codon CAG for residue 260 as Glu
C:Superfamily: actin
C:Keywords: methylated amino acid
F,7/3/Modified site: 3'-methylhistidine (His) #status predicted
```

Query Match 100.0%; Score 80; DB 2; Length 370;  
 Best Local Similarity 100.0%; Pred. No. 5.9e-06;  
 Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 APRAVFPISIVGRPRH 15  
 |||||  
 DB 26 APRAVFPISIVGRPRH 40

## RESULT 13

ATBOG

actin beta - bovine (tentative sequence)  
 C:Species: Bos primigenius taurus (cattle)  
 C>Date: 30-Sep-1991 #sequence\_revision 30-Sep-1991 #text\_change 09-Jul-2004  
 C:Accession: E14185; A39105; A02999; A14185  
 R:Vandekerckhove, J.; Weber, K.  
 Eur. J. Biochem. 90, 451-462, 1978  
 A:Title: Actin amino-acid sequences. Comparison of actins from calf thymus, bovine brain  
 A:Reference number: A14185; MUID:79045349; PMID:213279  
 A:Accession: E14185  
 A:Molecule type: protein  
 A:Residues: 1-374 <VAN>  
 A:CROSS-references: UNIPROT:P60712  
 A:Note: only peptides that differed in composition from the corresponding peptides of ra  
 R:Degeen, J.L.; Neubauer, M.G.; Degen, S.J.F.; Seyfried, C.E.; Morris, D.R.  
 J. Biol. Chem. 258, 12153-12162, 1983  
 A:Title: Regulation of protein synthesis in mitogen-activated bovine lymphocytes. Analys  
 A:Reference number: A39105; MUID:84032385; PMID:6195151  
 A:Accession: A39105  
 A:Molecule type: mRNA  
 A:Residues: 76-227;344-374 <DEG>  
 A:CROSS-references: GB:K00622; GB:K00623  
 A:Note: actins beta and gamma were not distinguished in this study  
 C:Comment: Vertebrate nonmuscle cells contain two highly conserved cytoskeletal actins,  
 C:Superfamily: actin  
 C:Keywords: blocked amino end; cell motility; cytoskeleton; methylated amino acid; micro  
 F:1/Modified site: blocked amino end (Asp) (probably acetylated) #status experimental  
 F:72/Modified site: 3'-methylhistidine (His) #status predicted

Query Match 100.0%; Score 80; DB 1; Length 374;  
 Best Local Similarity 100.0%; Pred. No. 5.9e-06;  
 Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 APRAVFPISIVGRPRH 15  
 |||||  
 DB 25 APRAVFPISIVGRPRH 39

## RESULT 14

ATBOG

actin gamma - bovine (tentative sequence)  
 C:Species: Bos primigenius taurus (cattle)  
 C>Date: 30-Sep-1991 #sequence\_revision 30-Sep-1991 #text\_change 09-Jul-2004  
 C:Accession: B14185; A02999  
 R:Vandekerckhove, J.; Weber, K.  
 Eur. J. Biochem. 90, 451-462, 1978  
 A:Title: Actin amino-acid sequences. Comparison of actins from calf thymus, bovine brain  
 A:Reference number: A14185; MUID:79045349; PMID:213279  
 A:Accession: B14185  
 A:Molecule type: protein  
 A:Residues: 1-374 <VAN>  
 A:CROSS-references: UNIPROT:P02571  
 A:Note: only peptides that differed in composition from the corresponding peptides of ra  
 C:Comment: Vertebrate nonmuscle cells contain two highly conserved cytoskeletal actins,  
 C:Superfamily: actin  
 C:Keywords: blocked amino end; cell motility; cytoskeleton; methylated amino acid; micro  
 F:1/Modified site: blocked amino end (Glu) (probably acetylated) #status experimental  
 F:72/Modified site: 3'-methylhistidine (His) #status predicted

Query Match 100.0%; Score 80; DB 1; Length 374;  
 Best Local Similarity 100.0%; Pred. No. 5.9e-06;  
 Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 APRAVFPISIVGRPRH 15  
 |||||  
 DB 25 APRAVFPISIVGRPRH 39

## RESULT 15

JC5818

gamma-actin - human  
 C:Species: Homo sapiens (man)  
 C>Date: 20-Feb-1998 #sequence\_revision 13-Mar-1998 #text\_change 09-Jul-2004  
 C:Accession: JC5818; PC4501  
 R:Hauschildt, S.; Schwarz, C.; Heine, H.; Ulmer, A.J.; Flad, H.D.; Rietschel, E.T.; Jens  
 Biochem. Biophys. Res. Commun. 241, 670-674, 1997  
 A:Title: Actin: A target of lipopolysaccharid-induced phosphorylation in human monocytes  
 A:Reference number: JC5818; MUID:98096379; PMID:9434766  
 A:Accession: JC5818  
 A:Molecule type: protein  
 A:Residues: 1-374 <HAU>  
 A:CROSS-references: UNIPROT:P02571  
 A:Experimental source: monocyte  
 A:Accession: PC4501  
 A:Molecule type: protein  
 A:Residues: 1-61;84-112;147-190;196-209;215-253;335-358 <HA2>  
 A:Experimental source: monocyte  
 C:Comment: This protein is involved in a signal transduction that eventually leads to mo  
 C:Superfamily: actin

Query Match 100.0%; Score 80; DB 2; Length 374;  
 Best Local Similarity 100.0%; Pred. No. 5.9e-06;  
 Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 APRAVFPISIVGRPRH 15  
 |||||  
 DB 25 APRAVFPISIVGRPRH 39

Search completed: April 8, 2005, 10:53:04  
 Job time : 10.3333 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: April 8, 2005, 09:58:56 ; Search time 57.8667 Seconds  
(without alignments)  
132.739 Million cell updates/sec

Title: US-09-423-351C-1

Perfect score: 80

Sequence: 1 APRAVFPSIVGRPRH 15

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1612378 seqs, 512079187 residues

Total number of hits satisfying chosen parameters: 1612378

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

UniProt\_03:\*

1: uniprot\_sprot:\*

2: uniprot\_trembl:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	80	100.0	20	2	Q8X127 magnaporthe
2	80	100.0	20	2	Q8X128 magnaporthe
3	80	100.0	20	2	Q95VV9 magnaporthe
4	80	100.0	20	2	Q95VM0 magnaporthe
5	80	100.0	20	2	Q96VW1 magnaporthe
6	80	100.0	20	2	Q96VW2 magnaporthe
7	80	100.0	20	2	Q9C2Z3 sclerotinia
8	80	100.0	20	2	Q9C2Z4 sclerotinia
9	80	100.0	20	2	Q9C2Z5 sclerotinia
10	80	100.0	20	2	Q9C2Z6 sclerotinia
11	80	100.0	20	2	Q9C2Z7 sclerotinia
12	80	100.0	20	2	Q9C2Z8 sclerotinia
13	80	100.0	20	2	Q9C2Z9 sclerotinia
14	80	100.0	20	2	Q9C300 sclerotinia
15	80	100.0	20	2	Q9C301 sclerotinia
16	80	100.0	20	2	Q9C302 sclerotinia
17	80	100.0	20	2	Q9C303 sclerotinia
18	80	100.0	20	2	Q9C304 sclerotinia
19	80	100.0	20	2	Q9C305 sclerotinia
20	80	100.0	20	2	Q9C306 sclerotinia
21	80	100.0	20	2	Q9C307 sclerotinia
22	80	100.0	20	2	Q9C308 sclerotinia
23	80	100.0	20	2	Q9C309 sclerotinia
24	80	100.0	20	2	Q9C310 sclerotinia
25	80	100.0	20	2	Q9C311 sclerotinia
26	80	100.0	20	2	Q9C312 sclerotinia
27	80	100.0	20	2	Q9C313 sclerotinia
28	80	100.0	20	2	Q9C314 sclerotinia
29	80	100.0	20	2	Q9C315 sclerotinia
30	80	100.0	20	2	Q9C316 sclerotinia
31	80	100.0	20	2	Q9C317 sclerotinia

32 80 100.0 20 2 Q9C318 sclerotinia  
33 80 100.0 20 2 Q9C319 sclerotinia  
34 80 100.0 20 2 Q9C320 sclerotinia  
35 80 100.0 20 2 Q9C321 sclerotinia  
36 80 100.0 20 2 Q9C322 sclerotinia  
37 80 100.0 20 2 Q9C323 sclerotinia  
38 80 100.0 20 2 Q9C324 sclerotinia  
39 80 100.0 20 2 Q9C325 sclerotinia  
40 80 100.0 20 2 Q9C326 sclerotinia  
41 80 100.0 20 2 Q9C327 sclerotinia  
42 80 100.0 20 2 Q9C328 sclerotinia  
43 80 100.0 20 2 Q9C329 sclerotinia  
44 80 100.0 20 2 Q9C330 sclerotinia  
45 80 100.0 20 2 Q9C331 sclerotinia

#### ALIGNMENTS

##### RESULT 1

Q8X127 PRELIMINARY; PRT; 20 AA.  
AC Q8X127;  
DT 01-MAR-2002 (Tremblrel. 20, Created)  
DT 01-MAR-2002 (Tremblrel. 20, Last sequence update)  
DT 01-JUN-2003 (Tremblrel. 24, Last annotation update)  
DE Actin (Fragment).  
OS Magnaporthe grisea (Rice blast fungus) (Pyricularia grisea).  
OC Eukaryota; Fungi; Ascomycota; Pezizomycotina; Sordariomycetes;  
OC Sordariomycetes incertae sedis; Magnaporthaceae; Magnaporthae.  
OX NCBI\_TaxID=148305;  
RN [1]  
RP SEQUENCE FROM N.A.  
RA Couch B.C., Kohn L.M.;  
RT "A multilocus gene genealogy concordant with host preference indicates segregation of a new species, Magnaporthe oryzae, from M. grisea.";  
RL Mycologia 94:683-693(2002).  
DR EMBL; AY063735; AAL59410.1;  
DR GO; GO:0015629; C:actin cytoskeleton; IEA.  
DR GO; GO:0005200; F:structural constituent of cytoskeleton; IEA.  
DR Pfam; PF00022; Actin; 1.  
FT NON\_TER 1  
FT NON\_TER 20  
SQ SEQUENCE 20 AA; 2141 MW; B06CCBB097FA4BD CRC64;

Query Match 100.0%; Score 80; DB 2; Length 20;

Best Local Similarity 100.0%; Pred. No. 8.2e-07; Mismatches 0; Indels 0; Gaps 0;  
Matches 15; Conservative 0;

QY 1 APRAVFPSIVGRPRH 15

Db 4 APRAVFPSIVGRPRH 18

##### RESULT 2

Q8X128 PRELIMINARY; PRT; 20 AA.  
AC Q8X128;  
DT 01-MAR-2002 (Tremblrel. 20, Created)  
DT 01-MAR-2002 (Tremblrel. 20, Last sequence update)  
DT 01-JUN-2003 (Tremblrel. 24, Last annotation update)  
DE Actin (Fragment).  
OS Magnaporthe grisea (Rice blast fungus) (Pyricularia grisea).  
OC Eukaryota; Fungi; Ascomycota; Pezizomycotina; Sordariomycetes;  
OC Sordariomycetes incertae sedis; Magnaporthaceae; Magnaporthae.  
OX NCBI\_TaxID=148305;  
RN [1]  
RP SEQUENCE FROM N.A.  
RA Couch B.C., Kohn L.M.;  
RT "A multilocus gene genealogy concordant with host preference indicates segregation of a new species, Magnaporthe oryzae, from M. grisea.";  
RL Mycologia 94:683-693(2002).  
DR EMBL; AY063734; AAL59409.1; -.

QY 1 APRAVFPSIVGRPRH 15  
Db 4 APRAVFPSIVGRPRH 18

```
DR GO; GO:0015629; C:actin cytoskeleton; IEA.
DR GO; GO:0005200; F:structural constituent of cytoskeleton; IEA.
DR Pfam; PF00022; Actin; 1.
FT NON_TER 1 20
SQ SEQUENCE 20 AA; 2141 MW; B06CCBB097FA4ABD CRC64;

Query Match 100.0%; Score 80; DB 2; Length 20;
Best Local Similarity 100.0%; Pred. No. 8.2e-07;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 APRAVFPSIVGRPRH 15
Db 4 APRAVFPSIVGRPRH 18

RESULT 3
Q96VV9 ID Q96VV9 PRELIMINARY; PRT; 20 AA.
AC Q96VV9;
DT 01-DEC-2001 (TrEMBLrel. 19, Created)
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE Actin (Fragment).
GN Name=act;
OS Magnaporthe salvinii.
OC Eukaryota; Fungi; Ascomycota; Pezizomycotina; Sordariomycetes;
OC Sordariomycetes incertae sedis; Magnaporthaceae; Magnaporthe.
OX NCBI_TaxID=165778;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=MS-1;
RA Couch B.C.; Kohn L.M.;
RT "A multilocus gene genealogy concordant with host preference indicates
RT segregation of a new species, Magnaporthe oryzae, from M. grisea.";
RL Mycologia 94:683-693(2002).
DR EMBL; AF395975; AAK77897.1; -.
DR GO; GO:0015629; C:actin cytoskeleton; IEA.
DR GO; GO:0005200; F:structural constituent of cytoskeleton; IEA.
DR InterPro; IPR004000; Actin_like.
DR Pfam; PF00022; Actin; 1.
FT NON_TER 1 20
SQ SEQUENCE 20 AA; 2141 MW; B06CCBB097FA4ABD CRC64;

Query Match 100.0%; Score 80; DB 2; Length 20;
Best Local Similarity 100.0%; Pred. No. 8.2e-07;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 APRAVFPSIVGRPRH 15
Db 4 APRAVFPSIVGRPRH 18

RESULT 4
Q96VW0 ID Q96VW0 PRELIMINARY; PRT; 20 AA.
AC Q96VW0;
DT 01-DEC-2001 (TrEMBLrel. 19, Created)
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE Actin (Fragment).
GN Name=act;
OS Magnaporthe rhizophila.
OC Eukaryota; Fungi; Ascomycota; Pezizomycotina; Sordariomycetes;
OC Sordariomycetes incertae sedis; Magnaporthaceae; Magnaporthe.
OX NCBI_TaxID=165777;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=ATCC 96043;
RA Couch B.C.; Kohn L.M.;
RT "A multilocus gene genealogy concordant with host preference indicates
RT segregation of a new species, Magnaporthe oryzae, from M. grisea.";
```

```
RL Mycologia 94:683-693(2002).
DR EMBL; AF395974; AAK77896.1; -.
DR GO; GO:0015629; C:actin cytoskeleton; IEA.
DR GO; GO:0005200; F:structural constituent of cytoskeleton; IEA.
DR InterPro; IPR004000; Actin_like.
DR Pfam; PF00022; Actin; 1.
FT NON_TER 1 20
SQ SEQUENCE 20 AA; 2141 MW; B06CCBB097FA4ABD CRC64;

Query Match 100.0%; Score 80; DB 2; Length 20;
Best Local Similarity 100.0%; Pred. No. 8.2e-07;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 APRAVFPSIVGRPRH 15
Db 4 APRAVFPSIVGRPRH 18

RESULT 5
Q96VW1 ID Q96VW1 PRELIMINARY; PRT; 20 AA.
AC Q96VW1;
DT 01-DEC-2001 (TrEMBLrel. 19, Created)
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE Actin (Fragment).
GN Name=act;
OS Magnaporthe poae (Kentucky bluegrass fungus).
OC Eukaryota; Fungi; Ascomycota; Pezizomycotina; Sordariomycetes;
OC Sordariomycetes incertae sedis; Magnaporthaceae; Magnaporthe.
OX NCBI_TaxID=148304;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=ATCC 6411;
RA Couch B.C.; Kohn L.M.;
RT "A multilocus gene genealogy concordant with host preference indicates
RT segregation of a new species, Magnaporthe oryzae, from M. grisea.";
RL Mycologia 94:683-693(2002).
DR EMBL; AF395973; AAK77895.1; -.
DR GO; GO:0015629; C:actin cytoskeleton; IEA.
DR GO; GO:0005200; F:structural constituent of cytoskeleton; IEA.
DR InterPro; IPR004000; Actin_like.
DR Pfam; PF00022; Actin; 1.
FT NON_TER 1 20
SQ SEQUENCE 20 AA; 2141 MW; B06CCBB097FA4ABD CRC64;

Query Match 100.0%; Score 80; DB 2; Length 20;
Best Local Similarity 100.0%; Pred. No. 8.2e-07;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 APRAVFPSIVGRPRH 15
Db 4 APRAVFPSIVGRPRH 18

RESULT 6
Q96VW2 ID Q96VW2 PRELIMINARY; PRT; 20 AA.
AC Q96VW2;
DT 01-DEC-2001 (TrEMBLrel. 19, Created)
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE Actin (Fragment).
GN Name=act;
OS Magnaporthe rhizophila.
OC Eukaryota; Fungi; Ascomycota; Pezizomycotina; Sordariomycetes;
OC Sordariomycetes incertae sedis; Magnaporthaceae; Magnaporthe.
OX NCBI_TaxID=165777;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=ATCC 58114;
```

```
RA Couch B.C., Kohn L.M.;
RT "A multilocus gene genealogy concordant with host preference indicates
RT segregation of a new species, Magnaporthe oryzae, from M. grisea.";
RL Mycologia 94:683-693(2002).
DR EMBL; AF395972; AAK77894.1; -.
DR GO; GO:0015629; C:actin cytoskeleton; IEA.
DR InterPro; IPR004000; F:structural constituent of cytoskeleton; IEA.
DR Pfam; PF00022; Actin; 1.
FT NON_TER 1
FT NON_TER 20
SQ SEQUENCE 20 AA; 2141 MW; B06CCBB097FA4ABD CRC64;

Query Match 100.0%; Score 80; DB 2; Length 20;
Best Local Similarity 100.0%; Pred. No. 8.2e-07;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 APRVFPSPVIGRPRH 15
DB 4 APRVFPSPVIGRPRH 18

RESULT 7
Q9C223 ID Q9C223 PRELIMINARY; PRT; 20 AA.
AC Q9C223;
DT 01-JUN-2001 (TREMBlrel. 17, Created)
DT 01-JUN-2001 (TREMBlrel. 17, Last sequence update)
DT 01-JUN-2003 (TREMBlrel. 24, Last annotation update)
DE Actin (Fragment).
GN Name=act;
OS Sclerotinia sclerotiorum.
OC Eukaryota; Fungi; Ascomycota; Pezizomycotina; Leotiomyces;
OC Helotiales; Sclerotiniaceae; Sclerotinia.
OX NCBI_TaxID=5180;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=W3-A3-5;
RX MEDLINE=21247026; PubMed=11348503;
RA Carbone I., Kohn L.M.;
RT "A microbial population-species interface: nested cladistic and
RT coalescent inference with multilocus data.";
RL Mol. Ecol. 10:947-964(2001).
DR EMBL; AF340964; AAK09783.1; -.
DR GO; GO:0015629; C:actin cytoskeleton; IEA.
DR InterPro; IPR004000; F:structural constituent of cytoskeleton; IEA.
DR Pfam; PF00022; Actin; 1.
FT NON_TER 1
FT NON_TER 1
SQ SEQUENCE 20 AA; 2141 MW; B06CCBB097FA4ABD CRC64;

Query Match 100.0%; Score 80; DB 2; Length 20;
Best Local Similarity 100.0%; Pred. No. 8.2e-07;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 APRVFPSPVIGRPRH 15
DB 4 APRVFPSPVIGRPRH 18

RESULT 8
Q9C224 ID Q9C224 PRELIMINARY; PRT; 20 AA.
AC Q9C224;
DT 01-JUN-2001 (TREMBlrel. 17, Created)
DT 01-JUN-2001 (TREMBlrel. 17, Last sequence update)
DT 01-JUN-2003 (TREMBlrel. 24, Last annotation update)
DE Actin (Fragment).
GN Name=act;
OS Sclerotinia sclerotiorum.
OC Eukaryota; Fungi; Ascomycota; Pezizomycotina; Leotiomyces;
OC Helotiales; Sclerotiniaceae; Sclerotinia.
OX NCBI_TaxID=5180;
```

```

RN SEQUENCE FROM N.A.
RP STRAIN=W3-A3-2;
RX MEDLINE=21247026; PubMed=11348503;
RA Carbone I., Kohn L.M.;
RT "A microbial population-species interface: nested cladistic and
RT coalescent inference with multilocus data.";
RL Mol. Ecol. 10:947-964(2001).
DR EMBL; AF340963; AAK09782.1; -.
DR GO; GO:0015629; C:actin cytoskeleton; IEA.
DR InterPro; IPR004000; F:structural constituent of cytoskeleton; IEA.
DR Pfam; PF00022; Actin; 1.
FT NON_TER 1
FT NON_TER 1
SQ SEQUENCE 20 AA; 2141 MW; B06CCBB097FA4ABD CRC64;

Query Match 100.0%; Score 80; DB 2; Length 20;
Best Local Similarity 100.0%; Pred. No. 8.2e-07;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 APRVFPSPVIGRPRH 15
DB 4 APRVFPSPVIGRPRH 18

RESULT 9
Q9C225 ID Q9C225 PRELIMINARY; PRT; 20 AA.
AC Q9C225;
DT 01-JUN-2001 (TREMBlrel. 17, Created)
DT 01-JUN-2001 (TREMBlrel. 17, Last sequence update)
DT 01-JUN-2003 (TREMBlrel. 24, Last annotation update)
DE Actin (Fragment).
GN Name=act;
OS Sclerotinia sclerotiorum.
OC Eukaryota; Fungi; Ascomycota; Pezizomycotina; Leotiomyces;
OC Helotiales; Sclerotiniaceae; Sclerotinia.
OX NCBI_TaxID=5180;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=W3-A1-5;
RX MEDLINE=21247026; PubMed=11348503;
RA Carbone I., Kohn L.M.;
RT "A microbial population-species interface: nested cladistic and
RT coalescent inference with multilocus data.";
RL Mol. Ecol. 10:947-964(2001).
DR EMBL; AF340962; AAK09781.1; -.
DR GO; GO:0015629; C:actin cytoskeleton; IEA.
DR InterPro; IPR004000; F:structural constituent of cytoskeleton; IEA.
DR Pfam; PF00022; Actin; 1.
FT NON_TER 1
FT NON_TER 1
SQ SEQUENCE 20 AA; 2141 MW; B06CCBB097FA4ABD CRC64;

Query Match 100.0%; Score 80; DB 2; Length 20;
Best Local Similarity 100.0%; Pred. No. 8.2e-07;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 APRVFPSPVIGRPRH 15
DB 4 APRVFPSPVIGRPRH 18

RESULT 10
Q9C226 ID Q9C226 PRELIMINARY; PRT; 20 AA.
AC Q9C226;
DT 01-JUN-2001 (TREMBlrel. 17, Created)
DT 01-JUN-2001 (TREMBlrel. 17, Last sequence update)
DT 01-JUN-2003 (TREMBlrel. 24, Last annotation update)
DE Actin (Fragment).
GN Name=act;
OS Sclerotinia sclerotiorum.
```

```

OC Eukaryota; Fungi; Ascomycota; Pezizomycotina; Leotiomycetes;
OC Helotiales; Sclerotiniaceae; Sclerotinia.
OX NCBI_TaxID=5180;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=W3-A1-3;
RX MEDLINE=21247026; PubMed=11348503;
RA Carbone I., Kohn L.M.;
RT "A microbial population-species interface: nested cladistic and
RT coalescent inference with multilocus data.";
RL Mol. Ecol. 10:947-964(2001).
DR EMBL; AF340961; AAK09780.1; -.
DR GO; GO:0015629; C:actin cytoskeleton; IEA.
DR GO; GO:0005200; F:structural constituent of cytoskeleton; IEA.
DR InterPro; IPR004000; Actin_like.
DR Pfam; PF00022; Actin; 1.
FT NON TER 1
SQ SEQUENCE 20 AA; 2141 MW; B06CCBB097FA4ABD CRC64;

Query Match 100.0%; Score 80; DB 2; Length 20;
Best Local Similarity 100.0%; Pred. No. 8.2e-07;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 APRVFPISIVGRPRH 15
DB 4 APRVFPISIVGRPRH 18

RESULT 11
Q9C227
ID Q9C227 PRELIMINARY; PRT; 20 AA.
AC Q9C227;
DT 01-JUN-2001 (TrEMBLrel. 17, Created)
DT 01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE Actin (Fragment).
GN Name=act;
OS Sclerotinia sclerotiorum.
OC Eukaryota; Fungi; Ascomycota; Pezizomycotina; Leotiomycetes;
OC Helotiales; Sclerotiniaceae; Sclerotinia.
OX NCBI_TaxID=5180;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=W2-A4-5;
RX MEDLINE=21247026; PubMed=11348503;
RA Carbone I., Kohn L.M.;
RT "A microbial population-species interface: nested cladistic and
RT coalescent inference with multilocus data.";
RL Mol. Ecol. 10:947-964(2001).
DR EMBL; AF340960; AAK09779.1; -.
DR GO; GO:0015629; C:actin cytoskeleton; IEA.
DR GO; GO:0005200; F:structural constituent of cytoskeleton; IEA.
DR InterPro; IPR004000; Actin_like.
DR Pfam; PF00022; Actin; 1.
FT NON TER 1
SQ SEQUENCE 20 AA; 2141 MW; B06CCBB097FA4ABD CRC64;

Query Match 100.0%; Score 80; DB 2; Length 20;
Best Local Similarity 100.0%; Pred. No. 8.2e-07;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 APRVFPISIVGRPRH 15
DB 4 APRVFPISIVGRPRH 18

RESULT 12
Q9C228
ID Q9C228 PRELIMINARY; PRT; 20 AA.
AC Q9C228;
DT 01-JUN-2001 (TrEMBLrel. 17, Created)
DT 01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)

```

```

DE Actin (Fragment).
GN Name=act;
OS Sclerotinia sclerotiorum.
OC Eukaryota; Fungi; Ascomycota; Pezizomycotina; Leotiomycetes;
OC Helotiales; Sclerotiniaceae; Sclerotinia.
OX NCBI_TaxID=5180;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=W2-A4-4;
RX MEDLINE=21247026; PubMed=11348503;
RA Carbone I., Kohn L.M.;
RT "A microbial population-species interface: nested cladistic and
RT coalescent inference with multilocus data.";
RL Mol. Ecol. 10:947-964(2001).
DR EMBL; AF340959; AAK09778.1; -.
DR GO; GO:0015629; C:actin cytoskeleton; IEA.
DR GO; GO:0005200; F:structural constituent of cytoskeleton; IEA.
DR InterPro; IPR004000; Actin_like.
DR Pfam; PF00022; Actin; 1.
FT NON TER 1
SQ SEQUENCE 20 AA; 2141 MW; B06CCBB097FA4ABD CRC64;

Query Match 100.0%; Score 80; DB 2; Length 20;
Best Local Similarity 100.0%; Pred. No. 8.2e-07;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 APRVFPISIVGRPRH 15
DB 4 APRVFPISIVGRPRH 18

RESULT 13
Q9C229
ID Q9C229 PRELIMINARY; PRT; 20 AA.
AC Q9C229;
DT 01-JUN-2001 (TrEMBLrel. 17, Created)
DT 01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE Actin (Fragment).
GN Name=act;
OS Sclerotinia sclerotiorum.
OC Eukaryota; Fungi; Ascomycota; Pezizomycotina; Leotiomycetes;
OC Helotiales; Sclerotiniaceae; Sclerotinia.
OX NCBI_TaxID=5180;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=W2-A3-5;
RX MEDLINE=21247026; PubMed=11348503;
RA Carbone I., Kohn L.M.;
RT "A microbial population-species interface: nested cladistic and
RT coalescent inference with multilocus data.";
RL Mol. Ecol. 10:947-964(2001).
DR EMBL; AF340958; AAK09777.1; -.
DR GO; GO:0015629; C:actin cytoskeleton; IEA.
DR GO; GO:0005200; F:structural constituent of cytoskeleton; IEA.
DR InterPro; IPR004000; Actin_like.
DR Pfam; PF00022; Actin; 1.
FT NON TER 1
SQ SEQUENCE 20 AA; 2141 MW; B06CCBB097FA4ABD CRC64;

Query Match 100.0%; Score 80; DB 2; Length 20;
Best Local Similarity 100.0%; Pred. No. 8.2e-07;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 APRVFPISIVGRPRH 15
DB 4 APRVFPISIVGRPRH 18

RESULT 14
Q9C300
ID Q9C300 PRELIMINARY; PRT; 20 AA.
AC Q9C300;

```



Job time : 58.8667 secs

DT 01-JUN-2001 (Tremblrel. 17, Created)  
 DT 01-JUN-2001 (Tremblrel. 17, Last sequence update)  
 DT 01-JUN-2003 (Tremblrel. 24, Last annotation update)  
 DE Actin (Fragment).  
 GN Name=act;  
 OS Sclerotinia sclerotiorum.  
 OC Eukaryota; Fungi; Ascomycota; Pezizomycotina; Leotiomyces;  
 OC Helotiales; Sclerotiniaceae; Sclerotinia.  
 OX NCBI\_TaxID=5180;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=W2-A3-4;  
 RX MEDLINE=21247026; PubMed=11348503;  
 RA Carbone I., Kohn L.M.;  
 RT "A microbial population-species interface: nested cladistic and  
 RL coalescent inference with multilocus data."; Mol. Ecol. 10:947-964 (2001).  
 DR EMBL; AF340957; AAK09776.1;  
 DR GO; GO:0015629; C:actin cytoskeleton; IEA.  
 DR GO; GO:0005200; F:structural constituent of cytoskeleton; IEA.  
 DR InterPro; IPR004000; Actin\_like.  
 DR Pfam; PF00022; Actin; 1.  
 FT NON\_TER 1  
 SQ SEQUENCE 20 AA; 2141 MW; B06CCBB097FA4ABD CRC64;

Query Match 100.0%; Score 80; DB 2; Length 20;  
 Best Local Similarity 100.0%; Pred. No. 8.2e-07;  
 Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 APRAVFPSIVGRPRH 15  
 Db |||||  
 4 APRAVFPSIVGRPRH 18

## RESULT 15

Q9C301 PRELIMINARY; PRT; 20 AA.  
 AC Q9C301;  
 DT 01-JUN-2001 (Tremblrel. 17, Created)  
 DT 01-JUN-2001 (Tremblrel. 17, Last sequence update)  
 DT 01-JUN-2003 (Tremblrel. 24, Last annotation update)  
 DE Actin (Fragment).  
 GN Name=act;  
 OS Sclerotinia sclerotiorum.  
 OC Eukaryota; Fungi; Ascomycota; Pezizomycotina; Leotiomyces;  
 OC Helotiales; Sclerotiniaceae; Sclerotinia.  
 OX NCBI\_TaxID=5180;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=W1-A5-4;  
 RX MEDLINE=21247026; PubMed=11348503;  
 RA Carbone I., Kohn L.M.;  
 RT "A microbial population-species interface: nested cladistic and  
 RL coalescent inference with multilocus data."; Mol. Ecol. 10:947-964 (2001).  
 DR EMBL; AF340956; AAK09775.1;  
 DR GO; GO:0015629; C:actin cytoskeleton; IEA.  
 DR GO; GO:0005200; F:structural constituent of cytoskeleton; IEA.  
 DR InterPro; IPR004000; Actin\_like.  
 DR Pfam; PF00022; Actin; 1.  
 FT NON\_TER 1  
 SQ SEQUENCE 20 AA; 2141 MW; B06CCBB097FA4ABD CRC64;

Query Match 100.0%; Score 80; DB 2; Length 20;  
 Best Local Similarity 100.0%; Pred. No. 8.2e-07;  
 Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 APRAVFPSIVGRPRH 15  
 Db |||||  
 4 APRAVFPSIVGRPRH 18

Search completed: April 8, 2005, 12:03:04

**THIS PAGE BLANK (USPTO)**

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

# OM protein - protein search, using sw model

Run on: April 8, 2005, 09:58:56 ; Search time 62.4 Seconds  
(without alignments)  
92.971 Million cell updates/sec

Title: US-09-423-351C-2

Perfect score: 80

Sequence: 1 FPSIVGRPRHQGVNV 15

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 2105692 seqs, 386760381 residues

Total number of hits satisfying chosen parameters: 2105692

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : A\_Geneseq\_16Dec04:\*

- 1: geneseqp1980s:\*
- 2: geneseqp1990s:\*
- 3: geneseqp2000s:\*
- 4: geneseqp2001s:\*
- 5: geneseqp2002s:\*
- 6: geneseqp2003as:\*
- 7: geneseqp2003bs:\*
- 8: geneseqp2004s:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	80	100.0	15	2 AAW92528	Beta-acti
2	80	100.0	105	5 ABP42469	Abp42469 Human ova
3	80	100.0	157	4 ABB66853	Abb66853 Drosophil
4	80	100.0	201	5 ABP43176	Abp43176 Human ova
5	80	100.0	204	5 ABP43016	Abp43016 Human ova
6	80	100.0	310	8 ADR38362	Adr38362 pigA3Cy3G
7	80	100.0	342	7 ADM05323	Adm05323 Human pro
8	80	100.0	374	3 AAB12985	Aab12985 Human bet
9	80	100.0	375	1 AAP61532	Aap61532 Sequence
10	80	100.0	375	2 AAR50328	Aar50328 Drug resi
11	80	100.0	375	3 AAB15017	Aab15017 Posttrans
12	80	100.0	375	3 AAY94569	Aay94569 Human car
13	80	100.0	375	3 AAB15016	Aab15016 Posttrans
14	80	100.0	375	5 ABB77395	Abb77395 Human act
15	80	100.0	375	6 ABR64371	AbR64371 Angiogene
16	80	100.0	375	6 ABR04830	AbR04830 Rat cytop
17	80	100.0	375	7 ADB85212	AdB85212 Rat actin
18	80	100.0	375	7 ADB61174	AdB61174 Rat prote
19	80	100.0	375	7 ADF30525	AdF30525 Rat angio
20	80	100.0	375	7 ADI63062	AdI63062 Human apo
21	80	100.0	375	7 ADI62970	AdI62970 Human apo
22	80	100.0	375	7 ADI63040	AdI63040 Human apo
23	80	100.0	375	8 ADL13002	AdL13002 Human sct
24	80	100.0	375	8 ADJ78489	AdJ78489 Actin pro
25	80	100.0	375	8 ADL14103	AdL14103 Human sar

26	80	100.0	375	8 ADP04899	AdP04899 Sea squir
27	80	100.0	375	8 ADP12391	AdP12391 Protein e
28	80	100.0	375	8 ADQ26098	AdQ26098 Gamma 1 a
29	80	100.0	375	8 ABO84772	ABO84772 Human can
30	80	100.0	375	8 ABM80841	ABm80841 Tumour-as
31	80	100.0	375	8 ADN23274	AdN23274 Bacterial
32	80	100.0	375	8 ADS88828	AdS88828 Amino aci
33	80	100.0	375	8 ADS88825	AdS88825 Amino aci
34	80	100.0	375	8 ADS88826	AdS88826 Amino aci
35	80	100.0	375	8 ADS88827	AdS88827 Amino aci
36	80	100.0	375	2 AAW19799	Aaw19799 Gamma-emo
37	80	100.0	376	4 ABB58162	Abb58162 Drosophil
38	80	100.0	376	4 ABB61322	Abb61322 Drosophil
39	80	100.0	376	4 ABB60354	Abb60354 Drosophil
40	80	100.0	376	4 ABB64853	Abb64853 Drosophil
41	80	100.0	376	4 ABB63315	Abb63315 Drosophil
42	80	100.0	376	6 ABR62327	AbR62327 Pacific w
43	80	100.0	376	8 ADN03845	Adn03845 Antipsori
44	80	100.0	377	3 AAB15014	Aab15014 Human car
45	80	100.0	377	3 AAY94568	Aay94568 Human car

## ALIGNMENTS

### RESULT 1

AAW92528

ID AAW92528 standard; peptide; 15 AA.

AC AAW92528;

DT 26-APR-1999 (first entry)

DE Beta-actin reference peptide substrate #2.

KW Peptide substrate; CCT; eukaryotic type II chaperonin complex; cyclin; binding agent; substrate-binding site; SBS; substrate folding; actin; tubulin; treatment; cancer; anticancer drug; viral infection; screening; reduced toxicity.

XX OS Synthetic.

XX WO9853322-A1.

PN PD

XX 26-NOV-1998.

PF 22-MAY-1998; 98WO-GB001485.

XX 23-MAY-1997; 97GB-00010762.

(CANC-) INST CANCER RES ROYAL CANCER HOSPITAL.

Willison K, Hynes G, Liou AK;

WPI; 1999-070162/06.

Identifying specific binding agents for substrate binding site in CCT chaperonin complex - also new peptide binding agents and their mimetics, and peptides containing a specific CCT binding site, used for treating cancer.

Disclosure; Fig 10; 97pp; English.

This invention describes a method which uses the CCT (eukaryotic type II chaperonin) complex or part of it, for identifying a binding agent that can occupy a substrate-binding site (SBS) on the CCT complex. By binding to the CCT complex, the binding agents block an SBS so that biological activity of the CCT complex is affected, particularly its ability to fold substrates such as actin, tubulin and cyclin. The binding agents are useful for treatment of cancer, particularly when used in combination with an anticancer drug, or viral infections. Nucleic acid fragments are used to screen for agents, e.g. binding agents that modulate interaction between the CCT complex and a protein that is to be folded. The binding

CC agents may target cells that are actively synthesizing tubulin etc.  
 CC (unlike known microtubule-stabilising agents that affect all cells), so  
 CC should have reduced toxicity for normal cells. AAM92527-W92541 are  
 CC peptide substrates used in the method of the invention  
 XX  
 SQ Sequence 15 AA;  
 Query Match 100.0%; Score 80; DB 2; Length 15;  
 Best Local Similarity 100.0%; Pred. No. 5.1e-07; Mismatches 0; Indels 0; Gaps 0;  
 Matches 15; Conservative 0;  
 XX  
 QY 1 FPSIVGRPRHQGVV 15  
 DB 1 FPSIVGRPRHQGVV 15  
 RESULT 2  
 ID ABP42469 standard; protein; 105 AA.  
 XX  
 AC ABP42469;  
 XX  
 DT 22-AUG-2002 (first entry)  
 XX  
 DE Human ovarian antigen HOCQ94, SEQ ID NO:3601.  
 XX  
 KW Human; ovarian antigen; ovary; ovarian; breast; cancer; tumour;  
 KW ovarian cancer; breast cancer; tumour; reproductive system disorder;  
 KW infertility; pregnancy disorder; anovulation; polycystic ovary syndrome;  
 KW PCOS; ovarian cyst; dysmenorrhoea; endocrine disorder; infection;  
 KW inflammatory condition; immune disorder; blood disorder;  
 KW cardiovascular disorder; respiratory disorder; neurological disorder;  
 KW gastrointestinal disorder; urinary system disorder; drug screening;  
 KW gene therapy; chromosome mapping; forensic analysis;  
 KW antibody preparation; cytostatic; immunomodulatory; neuroprotective;  
 KW antiinflammatory; gynaecological; reproductive.  
 XX  
 OS Homo sapiens.  
 XX  
 PN WO200200677-A1.  
 XX  
 PD 03-JAN-2002.  
 XX  
 PF 07-JUN-2001; 2001WO-US018569.  
 XX  
 PR 07-JUN-2000; 2000US-0209467P.  
 XX  
 PA (HUMA-) HUMAN GENOME SCI INC.  
 XX  
 PI Birse CE, Rosen CA;  
 XX  
 DR WPI, 2002-147878/19.  
 DR N-PSDB; ABQ55346.  
 XX  
 PT Isolated nucleic acid molecules encoding novel ovarian polypeptides,  
 PT useful in the prevention, treatment and diagnosis of cancer (e.g. ovarian  
 PT cancer), immune disorders, cardiovascular disorders and neurological  
 PT diseases.  
 XX  
 PS Claim 11; SEQ ID NO 3601; 2922pp; English.  
 XX  
 CC The invention relates to 2175 novel human ovarian antigens (ABP41054-  
 CC ABP43228) and to cDNAs encoding them (ABQ54131-ABQ56305), and also  
 CC encompasses polypeptides 90% identical and polynucleotides 95% identical  
 CC to the sequences of the invention. The invention additionally relates to  
 CC recombinant vectors and host cells comprising human ovarian antigen  
 CC polynucleotides, antibodies against human ovarian antigens, and the use  
 CC of ovarian antigen polynucleotides and polypeptides in diagnosing,  
 CC treating, prognosing or preventing various ovary and/or breast-related  
 CC disorders. Such conditions include ovarian cancer and breast cancer, and  
 CC metastatic tumours of ovarian or breast origin, reproductive system  
 CC disorders (e.g., infertility, disorders of pregnancy, anovulation,  
 CC polycystic ovary syndrome, ovarian cysts, and dysmenorrhoea), endocrine

CC disorders, infections (e.g., chlamydia, HIV, toxoplasmosis, and toxic  
 CC shock syndrome), inflammatory conditions (e.g., mastitis, oophoritis and  
 CC vaginitis), immune disorders (e.g., congenital and acquired  
 CC immunodeficiencies, autoimmune oophoritis, systemic lupus erythematosus),  
 CC blood-related disorders (e.g., anaemia), cardiovascular disorders,  
 CC respiratory disorders, neurological disorders, gastrointestinal disorders  
 CC and urinary system disorders. Ovarian antigen polypeptides and  
 CC polynucleotides may also be used in screening for compounds which  
 CC modulate ovarian antigen expression or activity. The polynucleotides may  
 CC further be used for gene therapy, chromosome mapping, in the  
 CC identification of individuals and in forensic analysis, and the  
 CC polypeptides may be used as food additives or to prepare antibodies  
 CC useful in disease diagnosis, drug targeting and phenotyping. The present  
 CC sequence represents a human ovarian antigen of the invention. Note: The  
 CC sequence data for this patent did not form part of the printed  
 CC specification, but was obtained in electronic format directly from WIPO  
 CC at ftp.wipo.int/pub/published\_pct\_sequences  
 XX  
 SQ Sequence 105 AA;  
 Query Match 100.0%; Score 80; DB 5; Length 105;  
 Best Local Similarity 100.0%; Pred. No. 3.9e-06;  
 Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 FPSIVGRPRHQGVV 15  
 DB 32 FPSIVGRPRHQGVV 46  
 RESULT 3  
 ID ABB66853 standard; protein; 157 AA.  
 XX  
 AC ABB66853;  
 XX  
 DT 26-MAR-2002 (first entry)  
 XX  
 DE Drosophila melanogaster polypeptide SEQ ID NO 27351.  
 XX  
 KW Drosophila; developmental biology; cell signalling; insecticide;  
 KW pharmaceutical.  
 XX  
 OS Drosophila melanogaster.  
 XX  
 PN WO200171042-A2.  
 XX  
 PD 27-SEP-2001.  
 XX  
 PF 23-MAR-2001; 2001WO-US009231.  
 XX  
 PR 23-MAR-2000; 2000US-0191637P.  
 PR 11-JUL-2000; 2000US-00614150.  
 XX  
 PA (PEKE ) PE CORP NY.  
 XX  
 PI Venter JC, Adams M, Li PWD, Myers EW;  
 XX  
 DR WPI, 2001-656860/75.  
 DR N-PSDB; ABL10956.  
 XX  
 PT New isolated nucleic acid detection reagent for detecting 1000 or more  
 PT genes from Drosophila and for elucidating cell signalling and cell-cell  
 PT interactions.  
 XX  
 PS Disclosure; SEQ ID NO 27351; 21pp + Sequence Listing; English.  
 XX  
 CC The invention relates to an isolated nucleic acid detection reagent  
 CC capable of detecting 1000 or more genes from Drosophila. The invention is  
 CC useful in developmental biology and in elucidating cell signalling and  
 CC cell-cell interactions in higher eukaryotes for the development of  
 CC insecticides, therapeutics and pharmaceutical drugs. The invention  
 CC discloses genomic DNA sequences (ABL16176-ABL30511), expressed DNA  
 CC sequences (ABL01840-ABL16175) and the encoded proteins (ABBS7737-

CC ABB72072). The sequence data for this patent did not form part of the  
CC printed specification, but was obtained in electronic format directly  
CC from WIPO at ftp.wipo.int/pub/published\_pct\_sequences  
XX  
SQ Sequence 157 AA;

Query Match 100.0%; Score 80; DB 4; Length 157;  
Best Local Similarity 100.0%; Pred. No. 5.9e-06;  
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 FPSIVGRPRHQGVWV 15  
| | | | | | | | | | | | | | | | | | | | | |  
DB 32 FPSIVGRPRHQGVWV 46

RESULT 4  
ABP43176  
ID ABP43176 standard; protein; 201 AA.  
XX AC  
XX ABP43176;  
XX  
XX 22-AUG-2002 (first entry)  
XX  
XX Human ovarian antigen HVQC49, SEQ ID NO:4308.  
XX  
XX Human; ovarian antigen; ovary; ovarian; breast; cancer; tumour;  
KW ovarian cancer; breast cancer; tumour; reproductive system disorder;  
KW infertility; pregnancy disorder; anovulation; polycystic ovary syndrome;  
KW PCOS; ovarian cyst; dysmenorrhea; endocrine disorder; infection;  
KW inflammatory condition; immune disorder; blood disorder;  
KW cardiovascular disorder; respiratory disorder; neurological disorder;  
KW gastrointestinal disorder; urinary system disorder; drug screening;  
KW gene therapy; chromosome mapping; forensic analysis;  
KW antibody preparation; cytostatic; immunomodulatory; neuroprotective;  
KW antiinflammatory; gynaecological; reproductive.  
XX  
XX Homo sapiens.  
XX OS  
XX WO200200677-A1.  
XX  
XX  
XX PD  
XX  
XX PF  
XX  
XX PR  
XX  
XX PA (HUMA-) HUMAN GENOME SCI INC.  
XX  
XX PI Birse CE, Rosen CA;  
XX  
XX WPI; 2002-147878/19.  
XX N-PSDB; ABQ56253.  
XX  
XX  
XX PT Isolated nucleic acid molecules encoding novel ovarian polypeptides,  
XX useful in the prevention, treatment and diagnosis of cancer (e.g. ovarian  
XX cancer), immune disorders, cardiovascular disorders and neurological  
XX diseases.  
XX  
XX  
XX PS Claim 11; SEQ ID NO 4308; 2922pp; English.  
XX  
XX  
XX The invention relates to 2175 novel human ovarian antigens (ABP41054-  
XX ABP43228) and to cDNAs encoding them (ABQ54131-ABQ56305), and also  
XX encompasses polypeptides 90% identical and polynucleotides 95% identical  
XX to the sequences of the invention. The invention additionally relates to  
XX recombinant vectors and host cells comprising human ovarian antigen  
XX polynucleotides, antibodies against human ovarian antigens, and the use  
XX of ovarian antigen polynucleotides and polypeptides in diagnosing,  
XX treating, prognosing or preventing various ovary and/or breast-related  
XX disorders. Such conditions include ovarian cancer and breast cancer, and  
XX metastatic tumours of ovarian or breast origin, reproductive system  
XX disorders (e.g., infertility, disorders of pregnancy, anovulation,  
XX polycystic ovary syndrome, ovarian cysts, and dysmenorrhoea), endocrine  
XX disorders, infections (e.g., chlamydia, HIV, toxoplasmosis) and toxic

shock syndrome), inflammatory conditions (e.g., mastitis, oophoritis and vaginitis), immune disorders (e.g., congenital and acquired immunodeficiencies, autoimmune oophoritis, systemic lupus erythematosus), blood-related disorders (e.g., anaemia), cardiovascular disorders, respiratory disorders, neurological disorders, gastrointestinal disorders and urinary system disorders. Ovarian antigen polypeptides and polynucleotides may also be used in screening for compounds which modulate ovarian antigen expression or activity. The polynucleotides may further be used for gene therapy, chromosome mapping, in the identification of individuals and in forensic analysis, and the polypeptides may be used as food additives or to prepare antibodies useful in disease diagnosis, drug targeting and phenotyping. The present sequence represents a human ovarian antigen of the invention. Note: The sequence data for this patent did not form part of the printed specification, but was obtained in electronic format directly from WIPO at [ftp.wipo.int/pub/published\\_pct\\_sequences](http://ftp.wipo.int/pub/published_pct_sequences)

Sequence 201 AA;

Query Match 100.0%; Score 80; DB 5; Length 201;  
Best Local Similarity 100.0%; Pred. No. 7.6e-06;  
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 FPSVIGRPRHQGMV 15  
||| |||||||  
DB 48 FPSVIGRPRHQGMV 62

RESULT 5  
ABP43016  
ID ABP43016 standard; protein; 204 AA.  
XX AC  
XX ABP43016;  
DT XX  
DE 22-AUG-2002 (first entry)  
XX DE Human ovarian antigen HSPSB24, SEQ ID NO:4148.  
XX OS Homo sapiens.  
XX OS WC200200677-A1.  
XX PN  
XX PD 03-JAN-2002.  
XX PF 07-JUN-2001; 2001WO-US018569.  
XX PR 07-JUN-2000; 2000US-0209467P.  
XX PA (HUMA-) HUMAN GENOME SCI INC.  
XX PI Birse CE, Rosen CA;  
XX DR WPI: 2002-147878/19.  
XX DR N-PSDB; ABQ56093.  
XX PT Isolated nucleic acid molecules encoding novel ovarian polypeptides,  
PT useful in the prevention, treatment and diagnosis of cancer (e.g. ovarian  
PT cancer), immune disorders, cardiovascular disorders and neurological  
PT diseases.  
XX PS Claim 11; SEQ ID NO 4148; 2922pp; English.  
XX PS

CC The invention relates to 2175 novel human ovarian antigens (ABP41054-  
 CC ABP43228) and to cDNAs encoding them (ABQ54131-ABQ56305), and also  
 CC to the sequences of the invention. The invention additionally relates to  
 CC recombinant vectors and host cells comprising human ovarian antigen  
 CC polynucleotides, antibodies against human ovarian antigens, and the use  
 CC of ovarian antigen polynucleotides and polypeptides in diagnosing,  
 CC treating, prognosing or preventing various ovary and/or breast-related  
 CC disorders. Such conditions include ovarian cancer and breast cancer, and  
 CC metastatic tumours of ovarian or breast origin, reproductive system  
 CC disorders (e.g., infertility, disorders of pregnancy, anovulation,  
 CC polycystic ovary syndrome, ovarian cysts, and dysmenorrhoea), endocrine  
 CC disorders, infections (e.g., chlamydia, HIV, toxoplasmosis, and toxic  
 CC shock syndrome), inflammatory conditions (e.g., mastitis, oophoritis and  
 CC vaginitis), immune disorders (e.g., congenital and acquired  
 CC immunodeficiencies, autoimmune oophoritis, systemic lupus erythematosus),  
 CC blood-related disorders (e.g., anaemia), cardiovascular disorders,  
 CC respiratory disorders, neurological disorders, gastrointestinal disorders  
 CC and urinary system disorders. Ovarian antigen polypeptides and  
 CC polynucleotides may also be used in screening for compounds which  
 CC modulate ovarian antigen expression or activity. The polynucleotides may  
 CC further be used for gene therapy, chromosome mapping, in the  
 CC identification of individuals and in forensic analysis, and the  
 CC polypeptides may be used as food additives or to prepare antibodies  
 CC useful in disease diagnosis, drug targeting and phenotyping. The present  
 CC sequence represents a human ovarian antigen of the invention. Note: The  
 CC sequence data for this patent did not form part of the printed  
 CC specification, but was obtained in electronic format directly from WIPO  
 CC at ftp.wipo.int/pub/published\_pct\_sequences  
 XX  
 SQ Sequence 204 AA;

Query Match 100.0%; Score 80; DB 5; Length 204;  
 Best Local Similarity 100.0%; Pred. No. 7.8e-06;  
 Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 PPSIVGRPRHQGVWV 15  
 |||||  
 DB 51 PPSIVGRPRHQGVWV 65

## RESULT 6

ADR38362  
 ID ADR38362 standard; protein; 310 AA.

AC ADR38362;

XX 18-NOV-2004 (first entry)

XX pigA3Cy3GFP plasmid protein product (Cycle3 GFP under an actin promoter).

XX piggyBac; transposon; transposase; gene transfer; plasmid; Cycle3 GFP;  
 KW actin promoter.

XX Bombyx mori.

OS Synthetic.

OS Unidentified.

XX Key Location/Qualifiers

FT Misc-difference 20 /note= "Encoded by GCC"

XX JP2004236642-A.

XX 26-AUG-2004.

XX 10-FEB-2003; 2003JP-00032306.

XX 10-FEB-2003; 2003JP-00032306.

XX (TORA ) TORAY IND INC.

XX WPI; 2004-608233/59.

DR N-PSDB; ADR38361.

XX Introducing foreign gene into insect cell, involves introducing DNA  
 PT having foreign gene integrated between inverted repeat sequences of  
 PT piggyBac transposon and piggyBac transposase that acts on inverted repeat  
 PT sequence.

XX Disclosure; Page; 25pp; Japanese.

XX This invention relates to a novel method for introducing a foreign gene  
 CC into an insect cell. Specifically, it refers to a foreign gene occurring  
 CC between inverted repeat sequences of a piggyBac transposon and using a  
 CC piggyBac transposase for gene transfer to give a high recombinant  
 CC efficiency. The present invention describes the insect cell as that of  
 CC Lepidoptera insect origin, preferably it is the silk worm Bombyx mori,  
 CC and stable integration occurs into the genome of this insect using the  
 CC transposase enzyme. In particular, this gene transfer method uses the  
 CC plasmid pigA3Cy3GFP which contains a Cycle3 GFP gene expressed under a  
 CC silk worm actin promoter (A3 promoter) between a pair of inverted repeat  
 CC sequences of the transposon piggyBac. This polypeptide sequence is the  
 CC protein product from the pigA3Cy3GFP plasmid DNA that represents the  
 CC silkworm actin promoter controlling expression of the Cycle3 GFP gene  
 CC given in an exemplification of the invention. NOTE: This sequence is  
 CC given as an embedded protein in the sequence listing and is not referred  
 CC to further within the specification.

XX Sequence 310 AA;

Query Match 100.0%; Score 80; DB 8; Length 310;  
 Best Local Similarity 100.0%; Pred. No. 1.2e-05;  
 Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 PPSIVGRPRHQGVWV 15

DB 32 PPSIVGRPRHQGVWV 46  
 |||||

## RESULT 7

ADM05323  
 ID ADM05323 standard; protein; 342 AA.

AC ADM05323;

XX 20-MAY-2004 (first entry)

XX Human protein of the invention SEQ ID NO:4008.

XX human; gene therapy; diagnostic marker; pharmaceutical.

XX Homo sapiens.

XX EPI347046-A1.

XX 24-SEP-2003.

XX 12-APR-2002; 2002EP-00008400.

XX 22-MAR-2002; 2002JP-00137785.

XX (REAS-) RES ASSOC BIOTECHNOLOGY.

XX Isogai T, Sugiyama T, Otsuki T, Wakamatsu A, Sato H, Ishii S;

XX Yamamoto J, Isono Y, Hio Y, Otsuka K, Nagai K, Irie R, Tamechika I;

XX Seki N, Yoshikawa T, Otsuka M, Nagahari K, Masuho Y;

XX WPI; 2003-723558/69.

XX N-PSDB; ADM02880.

XX New polynucleotides and polypeptides are useful in gene therapy, for  
 PT developing a diagnostic marker or medicines for regulating their  
 PT expression and activity, or as a target of gene therapy.

XX Claim 1; SEQ ID NO 4008; 305pp; English.

XX The invention relates to a novel human polynucleotide and the encoded  
 CC polypeptide. A polynucleotide of the invention may have a use in gene  
 CC therapy. An oligonucleotide of the invention ADM06202-ADM06773 is useful  
 CC as a primer for synthesizing the polynucleotide or as a probe for  
 CC detecting the polynucleotide. The polynucleotides ADM01316-ADM03758 are  
 CC useful in gene therapy, for developing a diagnostic marker or medicines  
 CC for regulating their expression and activity, or as a target of gene  
 CC therapy. The proteins ADM03759-ADM06201 encoded by the polynucleotides  
 CC are useful as pharmaceutical agents. The present sequence represents a  
 CC protein sequence of the invention.  
 XX Sequence 342 AA;  
 SQ

Query Match 100.0%; Score 80; DB 7; Length 342;  
 Best Local Similarity 100.0%; Pred. No. 1.3e-05;  
 Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 FPSIVGRPRHQGMV 15  
 |||||  
 DB 33 FPSIVGRPRHQGMV 47

RESULT 8  
 AAB12985  
 ID AAB12985 standard; protein; 374 AA.  
 XX  
 AC AAB12985;  
 XX  
 DT 29-NOV-2000 (first entry)  
 XX  
 DE Human beta-actin protein sequence.  
 XX  
 KW Beta actin; post translational modification; sickle cell anaemia;  
 KW irreversibly sickled cell; ISC; treatment.  
 XX  
 OS Homo sapiens.  
 XX  
 FN US6087398-A.  
 XX  
 PD 11-JUL-2000.  
 XX  
 PF 01-MAR-1996; 96US-00609236.  
 XX  
 PR 14-AUG-1995; 95US-0002288P.  
 XX  
 PA (SALA-) SOUTH ALABAMA MEDICAL SCI FOUND.  
 XX  
 PI Goodman SR;  
 XX  
 DR WPI; 2000-498210/44.  
 XX  
 PT Treatment for sickle cell anemia comprises administering a reducing agent  
 PT to inhibit and reverse sickled cell formation in blood.  
 XX  
 PS Example 17; Fig 5B; 53pp; English.  
 XX  
 CC This invention provides a treatment for sickle cell anaemia, which  
 CC comprises administering a reducing agent. The treatment inhibits the  
 CC formation of irreversibly sickled cells (ISC) and reverses ISC formation  
 CC in the blood. The present sequence represents the human beta-actin  
 CC protein. The sequence is used in the invention to demonstrate that a  
 CC disulphide bridge is formed between cysteines 284 and 373 in ISC beta-  
 CC actin as a post-translational modification. The reducing agent used in  
 CC the treatment acts to correct this post-translational modification  
 XX  
 SQ Sequence 374 AA;  
 Query Match 100.0%; Score 80; DB 3; Length 374;  
 Best Local Similarity 100.0%; Pred. No. 1.5e-05;  
 Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 FPSIVGRPRHQGMV 15

DB 30 FPSIVGRPRHQGMV 44  
 |||||  
 RESULT 9  
 AAP61532  
 ID AAP61532 standard; protein; 375 AA.  
 XX  
 AC AAP61532;  
 XX  
 DT 25-MAR-2003 (revised)  
 DT 13-SEP-1991 (first entry)  
 XX  
 DE Sequence of beta-actin.  
 XX  
 KW Actin fibre; actin film; recombinant beta-actin.  
 XX  
 OS Homo sapiens.  
 XX  
 PN EP174608-A.  
 XX  
 PD 19-MAR-1986.  
 XX  
 PF 05-SEP-1985; 85EP-00111225.  
 XX  
 PR 13-SEP-1984; 84US-00650958.  
 XX  
 PA (STRD ) UNIV LELAND STANFORD JUNIOR.  
 XX  
 PI Leavitt JC, Kedes LH, Gunning PW;  
 XX  
 DR WPI; 1986-077015/12.  
 DR N-PSDB; AAN60172.  
 XX  
 PT Beta-actin gene and regulatory elements - used for expression of  
 PT polypeptide(s) in mammalian host cells.  
 XX  
 PS Example; Page 23-24; 32pp; English.  
 XX  
 CC In the example, a beta-actin expression vector providing the beta- actin  
 CC promoter region, a polylinker and a polyadenylation signal was  
 CC constructed where the expression construct was present on a vector having  
 CC a bacterial origin of replication, as well as a marker for selection in a  
 CC mammalian host. (Updated on 25-MAR-2003 to correct PA field.)  
 XX  
 SQ Sequence 375 AA;  
 Query Match 100.0%; Score 80; DB 1; Length 375;  
 Best Local Similarity 100.0%; Pred. No. 1.5e-05;  
 Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 FPSIVGRPRHQGMV 15  
 |||||  
 DB 31 FPSIVGRPRHQGMV 45

RESULT 10  
 AAR50328  
 ID AAR50328 standard; protein; 375 AA.  
 XX  
 AC AAR50328;  
 XX  
 DT 12-OCT-1994 (first entry)  
 XX  
 DE Drug resistant structural protein.  
 XX  
 KW Drug resistant; structural gene; expression vector; selective; marker;  
 KW transformation; thymidine-kinase deleted cell.  
 XX  
 OS Homo sapiens.  
 XX  
 PN JP06038773-A.  
 XX

```

PD 15-FEB-1994.
XX
XX
PF 31-JAN-1992; 92JP-00045939.
XX
XX 31-JAN-1992; 92JP-00045939.
XX
XX (TOVM ) TOYOBO KK.
XX
XX WPI; 1994-094836/12.
XX N-PSDB; AAQ44861.
XX
XX Expression vector contg. drug resistant marker gene - useful for
XX transformation of thymidine kinase-deleted human cells for high levels of
XX foreign protein prodn.
XX
XX Disclosure; Page 5-7; 7pp; Japanese.
XX
XX This sequence is encoded by a drug resistant structural gene which may be
XX used in the expression vector of the invention. This gene is used as the
XX selective marker in the vector. The resulting vector may be used to
XX transform a thymidine-kinase deleted cell allowing introduction of a
XX foreign structural gene. The transformed cell may be used to produce
XX large amounts of useful protein
XX
XX Sequence 375 AA;
SQ
Query Match 100.0%; Score 80; DB 2; Length 375;
Best Local Similarity 100.0%; Pred. No. 1.5e-05;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 FPSIVGRPRHQGVNV 15
DB 31 FPSIVGRPRHQGVNV 45
|||||
|||||

RESULT 11
AAB15017
ID AAB15017 standard; protein; 375 AA.
XX
AC AAB15017;
XX
XX 07-DEC-2000 (first entry)
XX
XX Posttranslationally modified human cardiac actin mutant E361G.
XX
XX Cardiac actin; ACTC; human; gene therapy; IDC; chromosome 15q14;
XX idiopathic dilated cardiomyopathy; mutant; mutein.
XX
XX Homo sapiens.
XX Synthetic.
XX
XX Key Location/Qualifiers
FH Misc-difference 361
FT /note= "Wild-type Glu substituted by Gly"
XX
XX US6063576-A.
XX
XX 16-MAY-2000.
XX
XX 29-JUN-1998; 98US-00106217.
XX
XX 29-JUN-1998; 98US-00106217.
XX
XX (UTAH ) UNIV UTAH RES FOUND.
XX
XX Olson TM, Keating MT;
XX
XX WPI; 2000-375488/32.
XX N-PSDB; AAQ73739.
XX
XX New cardiac actin gene comprising histidine to arginine or glycine to
XX glutamic acid substitution, useful in the diagnosis of diseases
XX associated with the mutation, specifically idiopathic dilated

PT 15-FEB-1994.
XX
XX
PS Claim 3; Page; 36pp; English.
XX
XX The present invention relates to human cardiac actin (ACTC). Genotypic
XX analyses show that ACTC is linked to idiopathic dilated cardiomyopathy
XX (IDC). The ACTC gene maps to chromosome 15q14. Six PCR primer pairs
XX (AAA49020 to AAA49031) have been developed to allow analysis of the six
XX exons of ACTC. Studies of two families with IDC showed the presence of
XX two mutations, one in exon 5 and one in exon 6. The present sequence is a
XX mutant of posttranslationally modified ACTC without the initial two
XX residues. The mutation is a Glu to Gly mutation at position 361. The ACTC
XX protein may be used in a variety of methods for drug screening and for
XX rational drug design. The ACTC gene may be used to treat IDC by gene
XX therapy. Analysis of the ACTC gene provides early identification of
XX subjects likely to develop or who already have IDC. Note: The present
XX sequence is not shown in the specification but is derived from the ACTC
XX sequence described in AAY94569
XX
XX Sequence 375 AA;
SQ
Query Match 100.0%; Score 80; DB 3; Length 375;
Best Local Similarity 100.0%; Pred. No. 1.5e-05;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 FPSIVGRPRHQGVNV 15
DB 31 FPSIVGRPRHQGVNV 45
|||||
|||||

RESULT 12
AAY94569
ID AAY94569 standard; protein; 375 AA.
XX
AC AAY94569;
XX
XX 07-DEC-2000 (first entry)
XX
XX Human cardiac actin protein after posttranslational modification.
XX
XX Cardiac actin; ACTC; human; gene therapy; IDC; 15q14;
XX idiopathic dilated cardiomyopathy.
XX
XX Homo sapiens.
XX
XX US6063576-A.
XX
XX 16-MAY-2000.
XX
XX 29-JUN-1998; 98US-00106217.
XX
XX 29-JUN-1998; 98US-00106217.
XX
XX (UTAH ) UNIV UTAH RES FOUND.
XX
XX Olson TM, Keating MT;
XX
XX WPI; 2000-375488/32.
XX N-PSDB; AAA49032.
XX
XX New cardiac actin gene comprising histidine to arginine or glycine to
XX glutamic acid substitution, useful in the diagnosis of diseases
XX associated with the mutation, specifically idiopathic dilated
XX cardiomyopathy.
XX
XX Claim 3; Col 61-62; 36pp; English.
XX
XX The present sequence is human cardiac actin (ACTC) protein, without the
XX initial Met-Cys which is posttranslationally removed. Genotypic analyses
XX show that ACTC is linked to idiopathic dilated cardiomyopathy (IDC). The
XX ACTC gene maps to chromosome 15q14. Six PCR primer pairs (AAA49020 to
XX AAA49031) have been developed to allow analysis of the six exons of ACTC.
XX Studies of two families with IDC showed the presence of two mutations,

```



CC one in exon 5 and one in exon 6. The mutations are G to A in codon 312  
CC (Arg312His) and A to G in codon 361 (Glu361Gly). The ACTC protein may be  
CC used in a variety of methods for drug screening and for rational drug  
CC design. The ACTC gene may be used to treat IDC by gene therapy. Analysis  
CC of the ACTC gene provides early identification of subjects likely to  
CC develop or who already have IDC  
XX  
SQ Sequence 375 AA;

Query Match 100.0%; Score 80; DB 3; Length 375;  
Best Local Similarity 100.0%; Pred. No. 1.5e-05;  
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 FPSIVGRPRHQGMV 15  
DB 31 FPSIVGRPRHQGMV 45

RESULT 13  
AAB15016  
ID AAB15016 standard; protein; 375 AA.

AC AAB15016;

DT 07-DEC-2000 (first entry)

DE Posttranslationally modified human cardiac actin mutant R312H.

XX Cardiac actin; ACTC; human; gene therapy; IDC; chromosome 15q14;  
KW idiopathic dilated cardiomyopathy; mutant; mutein.

OS Homo sapiens.  
OS Synthetic.

FH Key Location/Qualifiers

FT Misc-difference 312  
FT /note= "Wild-type Arg substituted by His"

XX US6063576-A.

XX 16-MAY-2000.

PF 29-JUN-1998; 98US-00106217.

PR 29-JUN-1998; 98US-00106217.

XX (UTAH ) UNIV UTAH RES FOUND.

XX Olson TM, Keating MT;

XX WPI; 2000-375488/32.

DR N-PSDB; AAA73738.

XX New cardiac actin gene comprising histidine to arginine or glycine to  
PT glutamic acid substitution, useful in the diagnosis of diseases  
PT associated with the mutation, specifically idiopathic dilated  
PT cardiomyopathy.

PS Claim 3; Page; 36pp; English.

XX The present invention relates to human cardiac actin (ACTC). Genotypic  
CC analyses show that ACTC is linked to idiopathic dilated cardiomyopathy  
CC (IDC). The ACTC gene maps to chromosome 15q14. Six PCR primer pairs  
CC (AAA49020 to AAA49031) have been developed to allow analysis of the six  
CC exons of ACTC. Studies of two families with IDC showed the presence of  
CC two mutations, one in exon 5 and one in exon 6. The present sequence is a  
CC mutant of posttranslationally modified ACTC without the initial two  
CC residues. The mutation is a Arg to His mutation at position 312. The ACTC  
CC protein may be used in a variety of methods for drug screening and for  
CC rational drug design. The ACTC gene may be used to treat IDC by gene  
CC therapy. Analysis of the ACTC gene provides early identification of  
CC subjects likely to develop or who already have IDC. Note: The present  
CC sequence is not shown in the specification but is derived from the ACTC

CC sequence described in AAY94569

XX Sequence 375 AA;

Query Match 100.0%; Score 80; DB 3; Length 375;  
Best Local Similarity 100.0%; Pred. No. 1.5e-05;  
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 FPSIVGRPRHQGMV 15  
DB 31 FPSIVGRPRHQGMV 45

RESULT 14  
ABB77395  
ID ABB77395 standard; protein; 375 AA.

AC ABB77395;

DT 11-JUL-2002 (first entry)

DE Human actin gamma 1.

XX Human; dermatological; skin stress; ageing; spondin 2; cathepsin L;  
KW actin gamma 1; vimentin; fibroblast; skin; cosmetic; pharmaceutical.

OS Homo sapiens.

XX DE10050274-A1.

XX 18-APR-2002.

PF 09-OCT-2000; 2000DE-01050274.

PR 09-OCT-2000; 2000DE-01050274.

XX (HENK ) HENKEL KGAA.

XX Petersohn D, Schmitt G, Foerster T;

XX WPI; 2002-373046/41.

XX In vitro assays for skin stress and skin ageing includes determination of  
PT spondin 2, cathepsin L, actin gamma 1 and vimentin fragments secreted by  
PT skin fibroblasts.

PS Claim 6; Page 12; 14pp; German.

XX The invention relates to in vitro methods for the detection of skin  
CC stress and/or skin ageing in humans and animals based on the  
CC determination of spondin 2, cathepsin L, actin gamma 1 or vimentin  
CC fragments secreted by fibroblast from the skin under test. Use of the  
CC methods in a test for potential cosmetics and pharmaceuticals with an  
CC effect on these skin conditions and products containing vimentin  
CC fragments are also included. Products containing vimentin fragments are  
CC effective in the regulation, especially maintenance, of skin homeostasis

XX Sequence 375 AA;

Query Match 100.0%; Score 80; DB 5; Length 375;  
Best Local Similarity 100.0%; Pred. No. 1.5e-05;  
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 FPSIVGRPRHQGMV 15  
DB 31 FPSIVGRPRHQGMV 45

RESULT 15

ABR64271  
ID ABR64271 standard; protein; 375 AA.

XX ABR64271;

XX 15-OCT-2003 (first entry)  
 DT  
 XX  
 DE Angiogenesis protein BN0369.  
 XX  
 KW Cytostatic; antirheumatic; antiarthritic; antidiabetic; ophthalmological;  
 KW antipsoriatic; antiarteriosclerotic; cardiant; vasotropic; angiogenesis;  
 KW gene therapy; vasculature; cancer; rheumatoid arthritis; psoriasis;  
 KW diabetic retinopathy; cardiovascular disease; atherosclerosis;  
 KW ischemic limb disease; coronary artery disease.  
 XX  
 OS Homo sapiens.  
 XX  
 PN WO2003027285-A1.  
 XX  
 PD 03-APR-2003.  
 XX  
 XX 19-SEP-2002; 2002WO-AU001282.  
 XX  
 PR 27-SEP-2001; 2001AU-00007973.  
 PR 27-SEP-2001; 2001AU-00007974.  
 PR 11-OCT-2001; 2001AU-00008210.  
 PR 29-OCT-2001; 2001AU-00008532.  
 PR 13-NOV-2001; 2001AU-00008838.  
 PR 28-AUG-2002; 2002AU-00951032.  
 XX  
 PA (BION-) BIONOMICS LTD.  
 XX  
 PI Gamble JR, Hahn CN, Vadas MA;  
 XX  
 DR WPI; 2003-354655/33.  
 DR N-PSDB; ACF34548.  
 XX  
 PT New angiogenic genes and polypeptides, useful for diagnosing,  
 PT prognosticating or treating an angiogenesis-related disorder, e.g.  
 PT cancer, rheumatoid arthritis, diabetic retinopathy, psoriasis or  
 PT cardiovascular diseases.  
 XX  
 PS Claim 15; SEQ ID NO 206; 90pp; English.  
 XX  
 CC The invention relates to the isolation of novel genes (ACF34446-ACF34559)  
 CC encoding proteins (ABR64180-ABR64281) involved in the process of  
 CC angiogenesis. The nucleic acid molecules are useful in identifying and/or  
 CC obtaining full-length human genes involved in an angiogenic process. The  
 CC nucleic acid molecule, polypeptides or complexes encoded, cells or  
 CC genetically modified non-human animals derived from these are useful for  
 CC the screening of candidate pharmaceutical compounds used in treating  
 CC angiogenesis-related disorders. They are also useful for diagnosing,  
 CC prognosticating or treating an angiogenesis-related disorder, which  
 CC involves uncontrolled or enhanced angiogenesis or is a disorder in which  
 CC a decreased vasculature is of benefit (e.g. cancer, rheumatoid arthritis,  
 CC diabetic retinopathy, psoriasis or cardiovascular diseases such as  
 CC atherosclerosis), or involves inappropriately arrested or decreased  
 CC angiogenesis or is a disorder in which an expanding vasculature is of  
 CC benefit (e.g. ischemic limb disease or coronary artery disease). The  
 CC modulator of expression or activity of the polypeptide encoded by the  
 CC nucleic acid sequence is useful for manufacturing a medicament for the  
 CC treatment of an angiogenesis-related disorder. This sequence corresponds  
 CC to one of the novel angiogenic protein  
 XX  
 SQ Sequence 375 AA;  
 Query Match 100.0%; Score 80; DB 6; Length 375;  
 Best Local Similarity 100.0%; Pred. No. 1.5e-05;  
 Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 FPSIVGRPRHQGVV 15  
 Db 31 FPSIVGRPRHQGVV 45  

Search completed: April 8, 2005, 10:50:42  
 Job time : 64.4 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: April 8, 2005, 09:58:57 ; Search time 17.5333 Seconds  
(without alignments)  
63.863 Million cell updates/sec

Title: US-09-423-351C-2

Perfect score: 80

Sequence: 1 FPSIVGRPRHQGVWV 15

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 513545 seqs, 74649064 residues

Total number of hits satisfying chosen parameters: 513545

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Listing first 45 summaries

Database : Issued Patents AA:\*

- 1: /cgn2\_6/ptodata/1/1aa/5A COMB.pap:\*
- 2: /cgn2\_6/ptodata/1/1aa/5B COMB.pap:\*
- 3: /cgn2\_6/ptodata/1/1aa/6A COMB.pap:\*
- 4: /cgn2\_6/ptodata/1/1aa/6B COMB.pap:\*
- 5: /cgn2\_6/ptodata/1/1aa/PCTUS COMB.pap:\*
- 6: /cgn2\_6/ptodata/1/1aa/backfiles1.pap:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	80	100.0	374	3	US-08-609-236-6
2	80	100.0	375	2	US-08-494-151-14
3	80	100.0	375	3	US-09-106-217-16
4	80	100.0	375	4	US-09-976-594-731
5	80	100.0	376	1	US-08-588-113-2
6	80	100.0	376	4	US-09-949-016-6100
7	80	100.0	377	3	US-09-106-217-2
8	80	100.0	377	4	US-09-919-172-33
9	80	100.0	377	4	US-09-917-254-53
10	80	100.0	386	4	US-09-949-016-7721
11	80	100.0	399	4	US-09-949-016-9424
12	80	100.0	402	4	US-09-949-016-10757
13	80	100.0	404	4	US-09-949-016-11313
14	77	96.2	377	4	US-09-248-796A-14109
15	75	93.8	71	4	US-09-621-976-6521
16	74	92.5	20	3	US-08-505-250-17
17	74	92.5	20	3	US-08-505-250-17
18	74	92.5	374	3	US-09-306-446C-2
19	73	91.2	375	3	US-09-171-337A-7
20	73	91.2	375	3	US-09-171-337A-8
21	73	91.2	375	4	US-09-631-022-7
22	73	91.2	375	4	US-09-631-022-8
23	55	68.8	41	3	US-09-306-446C-3
24	52	65.0	376	4	US-09-538-092-1110
25	52	65.0	376	4	US-09-949-016-6656
26	52	65.0	376	4	US-09-949-016-8452
27	50	62.5	362	4	US-09-949-016-7725

28 50 62.5 376 4 US-09-538-092-1109 Sequence 1109, Ap  
29 48 60.0 11 4 US-09-786-066-5 Sequence 5, Appl  
30 47 58.8 504 4 US-09-554-726A-10 Sequence 10, Appl  
31 42 52.5 180 4 US-09-252-991A-31563 Sequence 11563, A  
32 41 51.2 85 4 US-09-248-796A-18932 Sequence 18932, A  
33 40 50.0 78 4 US-09-621-976-5181 Sequence 5181, Ap  
34 40 50.0 212 4 US-09-252-991A-29187 Sequence 29187, A  
35 39 48.8 150 4 US-09-732-210-626 Sequence 626, App  
36 39 48.8 374 4 US-09-252-991A-28830 Sequence 28830, A  
37 39 48.8 927 4 US-09-198-452A-472 Sequence 472, App  
38 39 48.8 937 4 US-09-438-185A-449 Sequence 449, App  
39 38 47.5 394 4 US-09-949-016-6655 Sequence 6655, Ap  
40 38 47.5 406 4 US-09-949-016-7396 Sequence 7396, Ap  
41 38 47.5 592 4 US-09-252-991A-22838 Sequence 22838, A  
42 38 47.5 630 4 US-09-252-991A-26324 Sequence 26324, A  
43 38 47.5 1220 1 US-08-158-232-43 Sequence 43, Appl  
44 38 47.5 1220 2 US-08-611-928-43 Sequence 43, Appl  
45 38 47.5 1220 3 US-09-173-891-43 Sequence 43, Appl

#### ALIGNMENTS

#### RESULT 1

US-08-609-236-6  
; Sequence 6, Application US/08609236  
; Patent No. 6087398  
; GENERAL INFORMATION:  
; APPLICANT: Steven R. Goodman  
; TITLE OF INVENTION: No. 6087398el Sickie Cell Anemia Treatment  
; NUMBER OF SEQUENCES: 6  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: McGregor & Adler, LLP  
; STREET: 8011 Candle Lane  
; CITY: Houston  
; STATE: Texas  
; COUNTRY: USA  
; ZIP: 77071  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: Apple Macintosh  
; OPERATING SYSTEM: Macintosh  
; SOFTWARE: Microsoft Word for Macintosh  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/609,236  
; FILING DATE: March 1, 1996  
; CLASSIFICATION: 514  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 60/002,288  
; FILING DATE: August 14, 1995  
; CLASSIFICATION: 514  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Benjamin Aaron Adler, Ph.D.  
; REGISTRATION NUMBER: 35,423  
; REFERENCE/DOCKET NUMBER: D5807  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 713-777-2321  
; TELEFAX: 713-777-6908  
; INFORMATION FOR SEQ ID NO: 6:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 374  
; TYPE: Amino acid  
; STRANDEDNESS:  
; TOPOLOGY: Linear  
; MOLECULE TYPE:  
; DESCRIPTION: Protein  
; HYPOTHETICAL: No  
; ANTI-SENSE: No  
; FRAGMENT TYPE:  
; ORIGINAL SOURCE:  
; STRAIN:  
; INDIVIDUAL ISOLATE:  
; DEVELOPMENTAL STAGE:

TISSUE TYPE:  
CELL TYPE:  
CELL LINE:

US-08-609-236-6

Query Match 100.0%; Score 80; DB 3; Length 374;  
Best Local Similarity 100.0%; Pred. No. 6.9e-06;  
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 FPSIVGRPRHQGVWV 15  
Db 30 FPSIVGRPRHQGVWV 44

## RESULT 2

US-08-494-151-14  
Sequence 14, Application US/08494151  
Patent No. 5840528  
GENERAL INFORMATION:  
APPLICANT: Van Ooyen, Albert Johannes Joseph  
TITLE OF INVENTION: Transformation of Phaffia rhodozyma  
NUMBER OF SEQUENCES: 14  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Morrison & Foerster  
STREET: 2000 Pennsylvania Ave. N.W., Suite 5500  
CITY: Washington, D.C.  
COUNTRY: USA  
ZIP: 20006-1812  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICANT: Morrison & Foerster  
FILING DATE: 23-JUN-1995  
CLASSIFICATION: 435  
ATTORNEY/AGENT INFORMATION:  
NAME: Murashige, Kate H.  
REGISTRATION NUMBER: 29,959  
REFERENCE/DOCKET NUMBER: 24615-20039.01  
TELEPHONE: (202) 887-1500  
TELEFAX: (202) 887-0763  
TELEX: 90-4030  
INFORMATION FOR SEQ ID NO: 14:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 375 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein

US-08-494-151-14

Query Match 100.0%; Score 80; DB 2; Length 375;  
Best Local Similarity 100.0%; Pred. No. 6.9e-06;  
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 FPSIVGRPRHQGVWV 15  
Db 31 FPSIVGRPRHQGVWV 45

## RESULT 3

US-09-106-217-16  
Sequence 16, Application US/09106217  
Patent No. 6063576  
GENERAL INFORMATION:  
APPLICANT: Keating, Mark T.  
APPLICANT: Olson, Timothy M.  
TITLE OF INVENTION: Actin Mutations in Dilated  
TITLE OF INVENTION: Cardiomyopathy, a Heritable Form of Heart Failure  
NUMBER OF SEQUENCES: 18  
CORRESPONDENCE ADDRESS:

ADDRESSEE: Rothwell, Figg, Ernst & Kurz, P.C.  
STREET: 555 Thirteenth Street, N.W., Suite 701 East  
CITY: Washington  
STATE: DC  
COUNTRY: U.S.A.  
ZIP: 20004

COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICANT: Morrison & Foerster  
FILING DATE: 23-JUN-1995  
CLASSIFICATION: 435  
ATTORNEY/AGENT INFORMATION:  
NAME: Saxe, Stephen A.  
REGISTRATION NUMBER: 38,609  
REFERENCE/DOCKET NUMBER: 2323-125  
TELEPHONE: 202-783-6040  
TELEFAX: 202-783-6031  
INFORMATION FOR SEQ ID NO: 16:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 375 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein

US-09-106-217-16

Query Match 100.0%; Score 80; DB 3; Length 375;  
Best Local Similarity 100.0%; Pred. No. 6.9e-06;  
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 FPSIVGRPRHQGVWV 15  
Db 31 FPSIVGRPRHQGVWV 45

## RESULT 4

US-09-976-594-731  
Sequence 731, Application US/09976594  
Patent No. 6673549  
GENERAL INFORMATION:  
APPLICANT: Furness, Michael  
APPLICANT: Buchbinder, Jenny  
TITLE OF INVENTION: GENES EXPRESSED IN C3A LIVER CELL CULTURES TREATED WITH STEROIDS  
FILE REFERENCE: PA-0041 US  
CURRENT APPLICATION NUMBER: US/09/976,594  
CURRENT FILING DATE: 2001-10-12  
PRIOR APPLICATION NUMBER: 60/240,409  
PRIOR FILING DATE: 2000-10-12  
NUMBER OF SEQ ID NOS: 1143  
SOFTWARE: PERL Program  
SEQ ID NO 731  
LENGTH: 375  
TYPE: PRT  
ORGANISM: Homo sapiens  
FEATURE:  
NAME/KEY: misc feature  
OTHER INFORMATION: Incyte ID No. 6673549 1837317CD1  
US-09-976-594-731

Query Match 100.0%; Score 80; DB 4; Length 375;  
Best Local Similarity 100.0%; Pred. No. 6.9e-06;  
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 FPSIVGRPRHQGVWV 15  
Db 31 FPSIVGRPRHQGVWV 45

## RESULT 5

US-08-588-113-2  
; Sequence 2, Application US/08588113  
; Patent No. 5710003  
; GENERAL INFORMATION:  
; APPLICANT: McHugh, Kirk M.  
; TITLE OF INVENTION: DIAGNOSTIC TEST FOR DETERMINING  
; TITLE OF INVENTION: MALIGNANCY OF SMOOTH MUSCLE TUMORS  
; NUMBER OF SEQUENCES: 4  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Woodcock Washburn Kurtz Mackiewicz &  
; ADDRESSEE: No. 5710003ris  
; STREET: One Liberty Place - 46th Floor  
; CITY: Philadelphia  
; STATE: PA  
; COUNTRY: USA  
; ZIP: 19103  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patent In Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/588,113  
; FILING DATE:  
; CLASSIFICATION: 436  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Ralph, Rebecca L.  
; REGISTRATION NUMBER: 35,152  
; REFERENCE/DOCKET NUMBER: TJU-1652  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 215-568-3100  
; TELEFAX: 215-568-3439  
; INFORMATION FOR SEQ ID NO: 2:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 376 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
US-08-588-113-2

Query Match 100.0%; Score 80; DB 1; Length 376;  
Best Local Similarity 100.0%; Pred. No. 6.9e-06;  
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 FPSIVGRPRHQGMV 15  
| | | | | | | | | | | | | | | | | | | | | |  
DB 32 FPSIVGRPRHQGMV 46

## RESULT 6

US-09-949-016-6100  
; Sequence 6100, Application US/09949016  
; Patent No. 6812339  
; GENERAL INFORMATION:  
; APPLICANT: VENTER, J. Craig et al.  
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED  
; TITLE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF  
; FILE REFERENCE: CL001307  
; CURRENT APPLICATION NUMBER: US/09/949,016  
; CURRENT FILING DATE: 2000-04-14  
; PRIOR APPLICATION NUMBER: 60/241,755  
; PRIOR FILING DATE: 2000-10-20  
; PRIOR APPLICATION NUMBER: 60/237,768  
; PRIOR FILING DATE: 2000-10-03  
; PRIOR APPLICATION NUMBER: 60/231,498  
; PRIOR FILING DATE: 2000-09-08  
; NUMBER OF SEQ ID NOS: 207012  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 6100  
; LENGTH: 376  
; TYPE: PRT  
; ORGANISM: Human

## US-09-949-016-6100

Query Match 100.0%; Score 80; DB 4; Length 376;  
Best Local Similarity 100.0%; Pred. No. 6.9e-06;  
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 FPSIVGRPRHQGMV 15  
| | | | | | | | | | | | | | | | | | | | | |  
DB 32 FPSIVGRPRHQGMV 46

## RESULT 7

US-09-106-217-2  
; Sequence 2, Application US/09106217  
; Patent No. 6063576  
; GENERAL INFORMATION:  
; APPLICANT: Keating, Mark T.  
; APPLICANT: Olson, Timothy M.  
; TITLE OF INVENTION: Actin Mutations in Dilated  
; TITLE OF INVENTION: Cardiomyopathy, a Heritable Form of Heart Failure  
; NUMBER OF SEQUENCES: 18  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Rothwell, Figg, Ernst & Kurz, P.C.  
; STREET: 555 Thirteenth Street, N.W., Suite 701 East  
; STREET: Tower  
; CITY: Washington  
; STATE: DC  
; COUNTRY: U.S.A.  
; ZIP: 20004  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patent In Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/09/106,217  
; FILING DATE:  
; CLASSIFICATION:  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Saxe, Stephen A.  
; REGISTRATION NUMBER: 38,609  
; REFERENCE/DOCKET NUMBER: 2323-125  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 202-783-6040  
; TELEFAX: 202-783-6031  
; INFORMATION FOR SEQ ID NO: 2:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 377 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
US-09-106-217-2

Query Match 100.0%; Score 80; DB 3; Length 377;  
Best Local Similarity 100.0%; Pred. No. 7e-06;  
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 FPSIVGRPRHQGMV 15  
| | | | | | | | | | | | | | | | | | | | | |  
DB 33 FPSIVGRPRHQGMV 47

## RESULT 8

US-09-919-172-33  
; Sequence 33, Application US/09919172  
; Patent No. 6673545  
; GENERAL INFORMATION:  
; APPLICANT: Paris, Mary  
; APPLICANT: Turner, Christopher M.  
; TITLE OF INVENTION: PROSTATE CANCER MARKERS  
; FILE REFERENCE: PA-0036 US  
; CURRENT APPLICATION NUMBER: US/09/919,172  
; CURRENT FILING DATE: 2001-07-30

```
; PRIOR APPLICATION NUMBER: 60/222,469
; PRIOR FILING DATE: 2000-07-28
; NUMBER OF SEQ ID NOS: 102
; SOFTWARE: PERL Program
; SEQ ID NO 33
; LENGTH: 377
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; OTHER INFORMATION: Incyte ID No. 6673545 17091118CD1
US-09-919-172-33

Query Match      100.0%; Score 80; DB 4; Length 377;
Best Local Similarity 100.0%; Pred. No. 7e-06;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      1 FPSIVGRPRHQGVV 15
Db      33 FPSIVGRPRHQGVV 47

RESULT 9
US-09-917-254-53
; Sequence 53, Application US/09917254
; Patent No. 6703204
; GENERAL INFORMATION:
; APPLICANT: Mutter, George
; TITLE OF INVENTION: Prognostic Classification of Breast Cancer
; FILE REFERENCE: B0801/7224(JRV)
; CURRENT APPLICATION NUMBER: US/09/917,254
; PRIOR FILING DATE: 2001-07-27
; PRIOR APPLICATION NUMBER: US 60/222,093
; PRIOR FILING DATE: 2000-07-28
; NUMBER OF SEQ ID NOS: 102
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 53
; LENGTH: 377
; TYPE: PRT
; ORGANISM: Homo Sapiens
US-09-917-254-53

Query Match      100.0%; Score 80; DB 4; Length 377;
Best Local Similarity 100.0%; Pred. No. 7e-06;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      1 FPSIVGRPRHQGVV 15
Db      33 FPSIVGRPRHQGVV 47

RESULT 10
US-09-949-016-7721
; Sequence 7721, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; PRIOR FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 7721
; LENGTH: 386
; TYPE: PRT
; ORGANISM: Human
US-09-949-016-7721

Query Match      100.0%; Score 80; DB 4; Length 402;
Best Local Similarity 100.0%; Pred. No. 7.4e-06;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      1 FPSIVGRPRHQGVV 15
Db      55 FPSIVGRPRHQGVV 69

RESULT 11
US-09-949-016-9424
; Sequence 9424, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-09-08
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 9424
; LENGTH: 399
; TYPE: PRT
; ORGANISM: Human
US-09-949-016-9424

Query Match      100.0%; Score 80; DB 4; Length 399;
Best Local Similarity 100.0%; Pred. No. 7.4e-06;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      1 FPSIVGRPRHQGVV 15
Db      55 FPSIVGRPRHQGVV 69

RESULT 12
US-09-949-016-10757
; Sequence 10757, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 10757
; LENGTH: 402
; TYPE: PRT
; ORGANISM: Human
US-09-949-016-10757

Query Match      100.0%; Score 80; DB 4; Length 402;
```

```

Best Local Similarity 100.0%; Pred. No. 7.5e-06;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 FPSIVGRPRHQGMV 15
Db 58 FPSIVGRPRHQGMV 72

RESULT 13
US-09-949-016-11313
; Sequence 11313, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; PRIOR FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 11313
; LENGTH: 404
; TYPE: PRT
; ORGANISM: Human
US-09-949-016-11313

Query Match 100.0%; Score 80; DB 4; Length 404;
Best Local Similarity 100.0%; Pred. No. 7.5e-06;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 FPSIVGRPRHQGMV 15
Db 60 FPSIVGRPRHQGMV 74

RESULT 14
US-09-248-796A-14109
; Sequence 14109, Application US/09248796A
; Patent No. 6747137
; GENERAL INFORMATION:
; APPLICANT: Keith Weinstock et al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO CANDIDA ALBICAN
; TITLE OF INVENTION: FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 107196.132
; CURRENT APPLICATION NUMBER: US/09/248,796A
; CURRENT FILING DATE: 1999-02-12
; PRIOR APPLICATION NUMBER: US 60/074,725
; PRIOR FILING DATE: 1998-02-13
; PRIOR APPLICATION NUMBER: US 60/096,409
; PRIOR FILING DATE: 1998-08-13
; NUMBER OF SEQ ID NOS: 28208
; SEQ ID NO 14109
; LENGTH: 377
; TYPE: PRT
; ORGANISM: Candida albicans
US-09-248-796A-14109

Query Match 96.2%; Score 77; DB 4; Length 377;
Best Local Similarity 86.7%; Pred. No. 2.3e-05;
Matches 13; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

Qy 1 FPSIVGRPRHQGMV 15
Db 33 FPSIVGRPRHQGMV 47

```

```

RESULT 15
US-09-621-976-6521
; Sequence 6521, Application US/09621976
; Patent No. 6639063
; GENERAL INFORMATION:
; APPLICANT: Dumas Milne Edwards, J.B.
; APPLICANT: Jobert, S.
; APPLICANT: Giordano, J.Y.
; TITLE OF INVENTION: ESTs and Encoded Human Proteins.
; FILE REFERENCE: GENSET.054PR2
; CURRENT APPLICATION NUMBER: US/09/621,976
; CURRENT FILING DATE: 2000-07-21
; NUMBER OF SEQ ID NOS: 19335
; SOFTWARE: Patent.pm
; SEQ ID NO 6521
; LENGTH: 71
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: UNSURE
; LOCATION: 62
; OTHER INFORMATION: Xaa = Ala,Ser
; NAME/KEY: UNSURE
; LOCATION: 43
; OTHER INFORMATION: Xaa = Leu,Val
US-09-621-976-6521

Query Match 93.8%; Score 75; DB 4; Length 71;
Best Local Similarity 93.3%; Pred. No. 8.4e-06;
Matches 14; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 FPSIVGRPRHQGMV 15
Db 31 FPSIVGRPRHQGMV 45

Search completed: April 8, 2005, 12:07:34
Job time : 18.5333 secs

```

**THIS PAGE BLANK (USPTO)**



GenCore version 5.1.6  
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: April 8, 2005, 10:53:18 ; Search time 42.6667 Seconds  
(without alignments)  
116.718 Million cell updates/sec

Title: US-09-423-351C-2

Perfect score: 80

Sequence: 1 FPSIVGRPRHOGVMV 15

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1418010 seqs, 331997259 residues

Total number of hits satisfying chosen parameters: 1418010

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

Published Applications AA:\*

- 1: /cgn2\_6/ptodata/1/pubpaa/US07\_PUBCOMB.pep:\*
- 2: /cgn2\_6/ptodata/1/pubpaa/PCT\_NEW\_PUB.pep:\*
- 3: /cgn2\_6/ptodata/1/pubpaa/US05\_NEW\_PUB.pep:\*
- 4: /cgn2\_6/ptodata/1/pubpaa/US06\_PUBCOMB.pep:\*
- 5: /cgn2\_6/ptodata/1/pubpaa/US07\_NEW\_PUB.pep:\*
- 6: /cgn2\_6/ptodata/1/pubpaa/PCTUS\_PUBCOMB.pep:\*
- 7: /cgn2\_6/ptodata/1/pubpaa/US08\_NEW\_PUB.pep:\*
- 8: /cgn2\_6/ptodata/1/pubpaa/US08\_PUBCOMB.pep:\*
- 9: /cgn2\_6/ptodata/1/pubpaa/US09A\_PUBCOMB.pep:\*
- 10: /cgn2\_6/ptodata/1/pubpaa/US09B\_PUBCOMB.pep:\*
- 11: /cgn2\_6/ptodata/1/pubpaa/US09C\_PUBCOMB.pep:\*
- 12: /cgn2\_6/ptodata/1/pubpaa/US09\_NEW\_PUB.pep:\*
- 13: /cgn2\_6/ptodata/1/pubpaa/US10A\_PUBCOMB.pep:\*
- 14: /cgn2\_6/ptodata/1/pubpaa/US10B\_PUBCOMB.pep:\*
- 15: /cgn2\_6/ptodata/1/pubpaa/US10C\_PUBCOMB.pep:\*
- 16: /cgn2\_6/ptodata/1/pubpaa/US10D\_PUBCOMB.pep:\*
- 17: /cgn2\_6/ptodata/1/pubpaa/US10\_NEW\_PUB.pep:\*
- 18: /cgn2\_6/ptodata/1/pubpaa/US11\_NEW\_PUB.pep:\*
- 19: /cgn2\_6/ptodata/1/pubpaa/US60\_NEW\_PUB.pep:\*
- 20: /cgn2\_6/ptodata/1/pubpaa/US60\_PUBCOMB.pep:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	80	100.0	105	15 US-10-264-049-3601	Sequence 3601, Ap
2	80	100.0	142	15 US-10-424-599-184471	Sequence 184471, Ap
3	80	100.0	201	15 US-10-264-049-4308	Sequence 4308, Ap
4	80	100.0	204	15 US-10-264-049-4148	Sequence 4148, Ap
5	80	100.0	219	14 US-10-108-631C-56	Sequence 56, Appl
6	80	100.0	342	15 US-10-108-260A-4008	Sequence 4008, Ap
7	80	100.0	371	16 US-10-322-281-68	Sequence 68, Appl
8	80	100.0	375	14 US-10-205-194-93	Sequence 93, Appl
9	80	100.0	375	14 US-10-316-253-88	Sequence 88, Appl
10	80	100.0	375	15 US-10-369-493-5927	Sequence 5927, Ap
11	80	100.0	375	15 US-10-205-331-94	Sequence 94, Appl
12	80	100.0	375	15 US-10-260-708-82	Sequence 82, Appl
13	80	100.0	376	15 US-10-341-434-101	Sequence 101, App

14	80	100.0	377	9 US-09-919-172-33	Sequence 33, Appl
15	80	100.0	377	15 US-10-236-031B-64	Sequence 64, Appl
16	80	100.0	377	16 US-10-333-680-4	Sequence 4, Appl
17	80	100.0	398	16 US-10-322-281-63	Sequence 63, Appl
18	80	100.0	413	9 US-09-925-301-1436	Sequence 1436, Ap
19	79	98.8	375	15 US-10-369-493-1786	Sequence 1786, Ap
20	79	98.8	375	16 US-10-477-369-1	Sequence 1, Appl
21	75	93.8	219	14 US-10-029-386-32056	Sequence 32056, A
22	74	92.5	135	16 US-10-767-701-47318	Sequence 47318, A
23	74	92.5	136	15 US-10-424-599-280811	Sequence 280811, A
24	74	92.5	197	15 US-10-424-599-280810	Sequence 280810, A
25	74	92.5	208	15 US-10-424-599-282905	Sequence 282905, A
26	74	92.5	230	16 US-10-767-701-42770	Sequence 42770, A
27	74	92.5	347	15 US-10-424-599-283333	Sequence 283333, A
28	74	92.5	375	15 US-10-369-493-2436	Sequence 2436, Ap
29	74	92.5	376	16 US-10-437-963-179917	Sequence 179917, A
30	74	92.5	376	16 US-10-437-963-202420	Sequence 202420, A
31	74	92.5	376	16 US-10-767-701-45677	Sequence 45677, A
32	74	92.5	376	16 US-10-767-701-46090	Sequence 46090, A
33	74	92.5	377	14 US-10-338-777-52	Sequence 52, Appl
34	74	92.5	377	15 US-10-424-599-162009	Sequence 162009, A
35	74	92.5	377	15 US-10-424-599-222816	Sequence 222816, A
36	74	92.5	377	15 US-10-424-599-223492	Sequence 223492, A
37	74	92.5	377	15 US-10-424-599-283336	Sequence 283336, A
38	74	92.5	377	15 US-10-425-114-66124	Sequence 66124, A
39	74	92.5	377	16 US-10-437-963-121952	Sequence 121952, A
40	74	92.5	377	16 US-10-437-963-148877	Sequence 148877, A
41	74	92.5	377	16 US-10-437-963-198295	Sequence 198295, A
42	74	92.5	377	16 US-10-767-701-47239	Sequence 47239, A
43	74	92.5	378	15 US-10-425-114-36824	Sequence 36824, A
44	74	92.5	378	15 US-10-425-114-42317	Sequence 42317, A
45	74	92.5	378	15 US-10-425-114-52458	Sequence 52458, A

## ALIGNMENTS

### RESULT 1

US-10-264-049-3601  
; Sequence 3601, Application US/10264049  
; Publication No. US20040005579A1  
; GENERAL INFORMATION:  
; APPLICANT: Birse et al.  
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies  
; FILE REFERENCE: PA133P1  
; CURRENT APPLICATION NUMBER: US/10/264,049  
; PRIOR FILING DATE: 2002-10-04  
; PRIOR APPLICATION NUMBER: PCT/US01/18569  
; PRIOR FILING DATE: 2001-06-07  
; PRIOR APPLICATION NUMBER: US 60/209,467  
; PRIOR FILING DATE: 2000-06-07  
; NUMBER OF SEQ ID NOS: 4360  
; SOFTWARE: PatentIn Ver. 3.1  
; SEQ ID NO 3601  
; LENGTH: 105  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
; FEATURE:  
; NAME/KEY: MISC FEATURE  
; LOCATION: (103)  
; OTHER INFORMATION: Xaa equals any of the twenty naturally occurring L-amino acids  
US-10-264-049-3601

Query Match 100.0%; Score 80; DB 15; Length 105;  
Best Local Similarity 100.0%; Pred. No. 2e-05;  
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 FPSIVGRPRHOGVMV 15

Db 32 FPSIVGRPRHOGVMV 46

### RESULT 2

```
US-10-424-599-184471
; Sequence 184471, Application US/10424599
; Publication No. US20040031072A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa Thomas J
; APPLICANT: Kovalic David K
; APPLICANT: Zhou Yihua
; APPLICANT: Cao Yongwei
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated with
; FILE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53223)B
; CURRENT APPLICATION NUMBER: US/10/424,599
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 285684
; SEQ ID NO 184471
; LENGTH: 142
; TYPE: PRT
; ORGANISM: Glycine max
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT3847_137594C.1.pep
US-10-424-599-184471

Query Match          100.0%; Score 80; DB 15; Length 142;
Best Local Similarity 100.0%; Pred. No. 2.7e-05;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 FPSIVGRPRHQGVNV 15
   |||||
Db 40 FPSIVGRPRHQGVNV 54

RESULT 3
US-10-264-049-4308
; Sequence 4308, Application US/10264049
; Publication No. US20040005579A1
; GENERAL INFORMATION:
; APPLICANT: Birse et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
; FILE REFERENCE: PA133P1
; CURRENT APPLICATION NUMBER: US/10/264,049
; CURRENT FILING DATE: 2002-10-04
; PRIOR APPLICATION NUMBER: PCT/US01/18569
; PRIOR FILING DATE: 2001-06-07
; PRIOR APPLICATION NUMBER: US 60/209,467
; PRIOR FILING DATE: 2000-06-07
; NUMBER OF SEQ ID NOS: 4360
; SOFTWARE: PatentIn Ver. 3.1
; SEQ ID NO 4308
; LENGTH: 201
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: MISC FEATURE
; LOCATION: (135)
; OTHER INFORMATION: Xaa equals any of the twenty naturally occurring L-amino acids
; NAME/KEY: MISC FEATURE
; LOCATION: (136)
; OTHER INFORMATION: Xaa equals any of the twenty naturally occurring L-amino acids
; NAME/KEY: MISC FEATURE
; LOCATION: (140)
; OTHER INFORMATION: Xaa equals any of the twenty naturally occurring L-amino acids
; FEATURE:
; NAME/KEY: MISC FEATURE
; LOCATION: (143)
; OTHER INFORMATION: Xaa equals any of the twenty naturally occurring L-amino acids
; NAME/KEY: MISC FEATURE
; LOCATION: (144)
; OTHER INFORMATION: Xaa equals any of the twenty naturally occurring L-amino acids
; NAME/KEY: MISC FEATURE
```

```
US-10-264-049-4308
; LOCATION: (145)
; OTHER INFORMATION: Xaa equals any of the twenty naturally occurring L-amino acids
; FEATURE:
; NAME/KEY: MISC FEATURE
; LOCATION: (151)
; OTHER INFORMATION: Xaa equals any of the twenty naturally occurring L-amino acids
; NAME/KEY: MISC FEATURE
; LOCATION: (159)
; OTHER INFORMATION: Xaa equals any of the twenty naturally occurring L-amino acids
; NAME/KEY: MISC FEATURE
; LOCATION: (172)
; OTHER INFORMATION: Xaa equals any of the twenty naturally occurring L-amino acids
; NAME/KEY: MISC FEATURE
; LOCATION: (174)
; OTHER INFORMATION: Xaa equals any of the twenty naturally occurring L-amino acids
; NAME/KEY: MISC FEATURE
; LOCATION: (180)
; OTHER INFORMATION: Xaa equals any of the twenty naturally occurring L-amino acids
US-10-264-049-4308

Query Match          100.0%; Score 80; DB 15; Length 201;
Best Local Similarity 100.0%; Pred. No. 3.9e-05;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 FPSIVGRPRHQGVNV 15
   |||||
Db 48 FPSIVGRPRHQGVNV 62

RESULT 4
US-10-264-049-4148
; Sequence 4148, Application US/10264049
; Publication No. US20040005579A1
; GENERAL INFORMATION:
; APPLICANT: Birse et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
; FILE REFERENCE: PA133P1
; CURRENT APPLICATION NUMBER: US/10/264,049
; CURRENT FILING DATE: 2002-10-04
; PRIOR APPLICATION NUMBER: PCT/US01/18569
; PRIOR FILING DATE: 2001-06-07
; PRIOR APPLICATION NUMBER: US 60/209,467
; PRIOR FILING DATE: 2000-06-07
; NUMBER OF SEQ ID NOS: 4360
; SOFTWARE: PatentIn Ver. 3.1
; SEQ ID NO 4148
; LENGTH: 204
; TYPE: PRT
; ORGANISM: Homo sapiens
; OTHER INFORMATION: Xaa equals any of the twenty naturally occurring L-amino acids
US-10-264-049-4148

Query Match          100.0%; Score 80; DB 15; Length 204;
Best Local Similarity 100.0%; Pred. No. 3.9e-05;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 FPSIVGRPRHQGVNV 15
   |||||
Db 51 FPSIVGRPRHQGVNV 65

RESULT 5
US-10-002-631C-56
; Sequence 56, Application US/10002631C
; Publication No. US20030157486A1
; GENERAL INFORMATION:
; APPLICANT: Graff, Jonathon M.
; APPLICANT: Muenster, Matthew
; TITLE OF INVENTION: METHODS TO IDENTIFY SIGNAL SEQUENCES
; FILE REFERENCE: A34943 090495.0243
```

; CURRENT APPLICATION NUMBER: US/10/002,631C  
; CURRENT FILING DATE: 2001-10-31  
; PRIOR APPLICATION NUMBER: 60/300,309  
; PRIOR FILING DATE: 2001-06-21  
; NUMBER OF SEQ ID NOS: 324  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 56  
; LENGTH: 219  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-10-002-631C-56

Query Match 100.0%; Score 80; DB 14; Length 219;  
Best Local Similarity 100.0%; Pred. No. 4.2e-05;  
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 FPSIVGRPRHOGVMV 15  
|||||

Db 58 FPSIVGRPRHOGVMV 72  
|||||

## RESULT 6

US-10-108-260A-4008  
; Sequence 4008, Application US/10108260A  
; Publication No. US20040005560A1  
; GENERAL INFORMATION:  
; APPLICANT: HELIX RESEARCH INSTITUTE  
; TITLE OF INVENTION: No. US20040005560A1el full length cDNA

; FILE REFERENCE: H1-A0106  
; CURRENT APPLICATION NUMBER: US/10/108,260A  
; CURRENT FILING DATE: 2002-03-27  
; NUMBER OF SEQ ID NOS: 5458  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 4008  
; LENGTH: 342  
; TYPE: PRT  
; ORGANISM: Homo sapiens

US-10-108-260A-4008

Query Match 100.0%; Score 80; DB 15; Length 342;  
Best Local Similarity 100.0%; Pred. No. 6.6e-05;  
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 FPSIVGRPRHOGVMV 15  
|||||

Db 33 FPSIVGRPRHOGVMV 47  
|||||

## RESULT 7

US-10-322-281-68  
; Sequence 68, Application US/10322281  
; Publication No. US20040126762A1  
; GENERAL INFORMATION:  
; APPLICANT: David W. Morris  
; TITLE OF INVENTION: Novel Compositions and Methods in Cancer  
; FILE REFERENCE: 529452001000  
; CURRENT APPLICATION NUMBER: US/10/322,281  
; CURRENT FILING DATE: 2002-12-17  
; NUMBER OF SEQ ID NOS: 866  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 68  
; LENGTH: 371  
; TYPE: PRT  
; ORGANISM: Homo sapiens

US-10-322-281-68

Query Match 100.0%; Score 80; DB 16; Length 371;  
Best Local Similarity 100.0%; Pred. No. 7.2e-05;  
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 FPSIVGRPRHOGVMV 15  
|||||

Db 31 FPSIVGRPRHOGVMV 45

## RESULT 8

US-10-205-194-93  
; Sequence 93, Application US/10205194  
; Publication No. US20030134301A1  
; GENERAL INFORMATION:  
; APPLICANT: Warner-Lambert Company  
; APPLICANT: Lee, Kevin  
; APPLICANT: Dixon, Alistair  
; APPLICANT: Brooksbank, Robert  
; APPLICANT: Pinnock, Robert  
; TITLE OF INVENTION: Identification and Use of Molecules Implicated in Pain  
; FILE REFERENCE: WL-A-018201  
; CURRENT APPLICATION NUMBER: US/10/205,194  
; CURRENT FILING DATE: 2000-07-24  
; PRIOR APPLICATION NUMBER: GB 0118354.0  
; PRIOR FILING DATE: 2001-07-27  
; NUMBER OF SEQ ID NOS: 177  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 93  
; LENGTH: 375  
; TYPE: PRT  
; ORGANISM: Rattus rattus  
; FEATURE:  
; OTHER INFORMATION: Cytoplasmic gamma isoform of actin

US-10-205-194-93

Query Match 100.0%; Score 80; DB 14; Length 375;  
Best Local Similarity 100.0%; Pred. No. 7.2e-05;  
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 FPSIVGRPRHOGVMV 15  
|||||

Db 31 FPSIVGRPRHOGVMV 45  
|||||

## RESULT 9

US-10-316-253-88  
; Sequence 88, Application US/10316253  
; Publication No. US20030162706A1  
; GENERAL INFORMATION:  
; APPLICANT: The Procter & Gamble Company  
; APPLICANT: Peters, Kevin  
; APPLICANT: Thompson, Larry  
; APPLICANT: Wang, Feng  
; APPLICANT: Greis, Kenneth  
; TITLE OF INVENTION: Angiogenesis Modulating Proteins  
; FILE REFERENCE: 8865M  
; CURRENT APPLICATION NUMBER: US/10/316,253  
; CURRENT FILING DATE: 2002-12-10  
; PRIOR APPLICATION NUMBER: US 60/355,295  
; PRIOR FILING DATE: 2002-02-08  
; NUMBER OF SEQ ID NOS: 308  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 88  
; LENGTH: 375  
; TYPE: PRT  
; ORGANISM: Rattus norvegicus

US-10-316-253-88

Query Match 100.0%; Score 80; DB 14; Length 375;  
Best Local Similarity 100.0%; Pred. No. 7.2e-05;  
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 FPSIVGRPRHOGVMV 15  
|||||

Db 31 FPSIVGRPRHOGVMV 45  
|||||

## RESULT 10

US-10-369-493-5927

```
; Sequence 5927, Application US/10369493
; Publication No. US20030233675A1
; GENERAL INFORMATION:
; APPLICANT: Cao, Yongwei
; APPLICANT: Hinkle, Gregory J.
; APPLICANT: Slater, Steven C.
; APPLICANT: Goldman, Barry S.
; APPLICANT: Chen, Xianfeng
; TITLE OF INVENTION: EXPRESSION OF MICROBIAL PROTEINS IN PLANTS FOR PRODUCTION OF
; FILE REFERENCE: 38-10(52052)B
; CURRENT APPLICATION NUMBER: US/10/369,493
; PRIOR APPLICATION NUMBER: US 60/360,039
; PRIOR FILING DATE: 2002-02-21
; NUMBER OF SEQ ID NOS: 47374
; SEQ ID NO 5927
; LENGTH: 375
; TYPE: PRT
; ORGANISM: Caenorhabditis elegans
US-10-369-493-5927

Query Match 100.0%; Score 80; DB 15; Length 375;
Best Local Similarity 100.0%; Pred. No. 7.2e-05;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 FPSIVGRPRHQGVWV 15
Db 31 FPSIVGRPRHQGVWV 45

RESULT 11
US-10-205-331-94
; Sequence 94, Application US/10205331
; Publication No. US20040058326A1
; GENERAL INFORMATION:
; APPLICANT: Warner-Lambert Company
; APPLICANT: Lee, Kevin
; APPLICANT: Dixon, Alistair
; APPLICANT: Brooksbank, Robert
; APPLICANT: Pinnock, Robert
; TITLE OF INVENTION: Identification and Use of Molecules Implicated in Pain
; FILE REFERENCE: WL-A-018199
; CURRENT APPLICATION NUMBER: US/10/205,331
; CURRENT FILING DATE: 2002-07-24
; PRIOR APPLICATION NUMBER: GB 0118354.0
; PRIOR FILING DATE: 2001-07-27
; NUMBER OF SEQ ID NOS: 117
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 94
; LENGTH: 375
; TYPE: PRT
; ORGANISM: Rattus norvegicus
; FEATURE:
; OTHER INFORMATION: Cytoplasmic beta-actin
US-10-205-331-94

Query Match 100.0%; Score 80; DB 15; Length 375;
Best Local Similarity 100.0%; Pred. No. 7.2e-05;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 FPSIVGRPRHQGVWV 15
Db 31 FPSIVGRPRHQGVWV 45

RESULT 12
US-10-260-708-82
; Sequence 82, Application US/10260708
; Publication No. US20040063101A1
; GENERAL INFORMATION:
; APPLICANT: Scanlan, Matthew
; APPLICANT: Lee, Sang-Yull
```

```
; APPLICANT: Old, Lloyd
; TITLE OF INVENTION: Human Sarcoma-Associated Antigens
; FILE REFERENCE: L00461/70138
; CURRENT APPLICATION NUMBER: US/10/260,708
; CURRENT FILING DATE: 2002-09-30
; NUMBER OF SEQ ID NOS: 96
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 82
; LENGTH: 375
; TYPE: PRT
; ORGANISM: homo sapiens
US-10-260-708-82

Query Match 100.0%; Score 80; DB 15; Length 375;
Best Local Similarity 100.0%; Pred. No. 7.2e-05;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 FPSIVGRPRHQGVWV 15
Db 31 FPSIVGRPRHQGVWV 45

RESULT 13
US-10-341-434-101
; Sequence 101, Application US/10341434
; Publication No. US20030215835A1
; GENERAL INFORMATION:
; APPLICANT: Origene Technologies
; TITLE OF INVENTION: Differentially Regulated Prostate Cancer Genes
; FILE REFERENCE: 9U 204 205 R1
; CURRENT APPLICATION NUMBER: US/10/341,434
; CURRENT FILING DATE: 2003-07-18
; PRIOR APPLICATION NUMBER: US 60/348,164
; PRIOR FILING DATE: 2002-01-15
; PRIOR APPLICATION NUMBER: US 60/348,119
; PRIOR FILING DATE: 2002-01-15
; NUMBER OF SEQ ID NOS: 238
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 101
; LENGTH: 376
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-341-434-101

Query Match 100.0%; Score 80; DB 15; Length 376;
Best Local Similarity 100.0%; Pred. No. 7.3e-05;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 FPSIVGRPRHQGVWV 15
Db 32 FPSIVGRPRHQGVWV 46

RESULT 14
US-09-919-172-33
; Sequence 33, Application US/09919172
; Patent No. US20020119463A1
; GENERAL INFORMATION:
; APPLICANT: Paris, Mary
; APPLICANT: Turner, Christopher M.
; TITLE OF INVENTION: PROSTATE CANCER MARKERS
; FILE REFERENCE: PA-0036 US
; CURRENT APPLICATION NUMBER: US/09/919,172
; CURRENT FILING DATE: 2001-07-30
; PRIOR APPLICATION NUMBER: 60/222,469
; PRIOR FILING DATE: 2000-07-28
; NUMBER OF SEQ ID NOS: 102
; SOFTWARE: PERL Program
; SEQ ID NO 33
; LENGTH: 377
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
```

; NAME/KEY: misc feature  
; OTHER INFORMATION: Incyte ID No. US20020119463A1 1709118CD1  
US-09-919-172-33

Query Match 100.0%; Score 80; DB 9; Length 377;  
Best Local Similarity 100.0%; Pred. No. 7.3e-05;  
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 FPSIVGRPRHQGMV 15  
| | | | | | | | | | | | | | |  
DB 33 FPSIVGRPRHQGMV 47

RESULT 15  
US-10-236-031B-64  
; Sequence 64, Application US/10236031B  
; Publication No. US20030219760A1  
; GENERAL INFORMATION:  
; APPLICANT: Gordon, Gavin J.  
; APPLICANT: Jensen, Roderick V.  
; APPLICANT: Gullans, Steven R.  
; APPLICANT: Bueno, Raphael  
; TITLE OF INVENTION: Diagnostic and Prognostic Tests  
; FILE REFERENCE: B00801/70265 (JRV/JAV)  
; CURRENT APPLICATION NUMBER: US/10/236,031B  
; CURRENT FILING DATE: 2002-09-05  
; PRIOR APPLICATION NUMBER: US 60/317,389  
; PRIOR FILING DATE: 2001-09-05  
; PRIOR APPLICATION NUMBER: US 60/407,431  
; PRIOR FILING DATE: 2002-08-30  
; NUMBER OF SEQ ID NOS: 102  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 64  
; LENGTH: 377  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-10-236-031B-64

Query Match 100.0%; Score 80; DB 15; Length 377;  
Best Local Similarity 100.0%; Pred. No. 7.3e-05;  
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 FPSIVGRPRHQGMV 15  
| | | | | | | | | | | | | | |  
DB 33 FPSIVGRPRHQGMV 47

Search completed: April 8, 2005, 12:50:57  
Job time : 42.6667 secs

**THIS PAGE BLANK (USPTO)**

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: April 8, 2005, 09:58:56 ; Search time 9.33333 Seconds  
(without alignments)  
154.634 Million cell updates/sec

Title: US-09-423-351C-2  
Perfect score.: 80  
Sequence: 1 FPSIVGRPRHOGVMV 15

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 283416 seqs, 96216763 residues  
Total number of hits satisfying chosen parameters: 283416

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000  
Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : PIR 79:.\*  
1: pir1.\*  
2: pir2.\*  
3: pir3.\*  
4: pir4.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	80	100.0	86	2	actin beta, cyto
2	80	100.0	137	2	actin 5C - fruit f
3	80	100.0	158	2	alpha-cardiac acti
4	80	100.0	213	2	actin CAL5 - sea s
5	80	100.0	308	2	actin 3 - fruit fl
6	80	100.0	328	2	actin beta - grass
7	80	100.0	349	2	actin, fetal skele
8	80	100.0	362	2	actin type 5, cyto
9	80	100.0	374	1	actin beta - bovin
10	80	100.0	374	1	actin gamma - bovi
11	80	100.0	374	2	gamma-actin - huma
12	80	100.0	375	1	actin, aortic smoo
13	80	100.0	375	1	actin, skeletal mu
14	80	100.0	375	1	actin beta - rat
15	80	100.0	375	1	actin beta, cytosk
16	80	100.0	375	1	actin beta - chick
17	80	100.0	375	1	actin beta - human
18	80	100.0	375	1	actin gamma 1 - hu
19	80	100.0	375	1	actin beta - mouse
20	80	100.0	375	1	actin gamma - mous
21	80	100.0	375	1	actin beta, non-mu
22	80	100.0	375	1	actin gamma, cytos
23	80	100.0	375	2	actin, cytosolic (
24	80	100.0	375	2	hypothetical prote
25	80	100.0	375	2	actin beta-2, cyto
26	80	100.0	375	2	actin beta-1, cyto
27	80	100.0	375	2	actin beta, cyto
28	80	100.0	375	2	actin beta - goose
29	80	100.0	375	2	actin alpha, cardi

30	80	100.0	375	2	S70377	actin - Phaffia rh
31	80	100.0	376	1	ATFF7	actin 7 - fruit fl
32	80	100.0	376	1	A43552	actin gamma, cytos
33	80	100.0	376	1	ATCHSM	actin gamma, smoot
34	80	100.0	376	1	ATFF8	actin 8 - fruit fl
35	80	100.0	376	1	ATURS	actin Cyl - sea ur
36	80	100.0	376	2	A48449	actin-1A - nematod
37	80	100.0	376	2	S04538	actin 87E - fruit
38	80	100.0	376	2	JC1246	actin - fruit fly
39	80	100.0	376	2	J50189	actin, cytosolic -
40	80	100.0	376	2	J50190	actin, muscle - st
41	80	100.0	376	2	S07288	actin 15A - sea ur
42	80	100.0	376	2	S09578	actin - sea urchin
43	80	100.0	376	2	A40261	actin gamma, enter
44	80	100.0	376	2	JQ0154	actin - Hydra atte
45	80	100.0	376	2	JN0832	actin (clone gen3)

## ALIGNMENTS

### RESULT 1

C43616  
actin beta, cytosolic - chicken (fragment)  
C:Species: Gallus gallus (chicken)  
C:Date: 11-Dec-1992 #sequence\_revision 11-Dec-1992 #text\_change 09-Jul-2004  
C:Accession: C43616  
R:Paterson, B.M.; Eldridge, J.D.  
Science 224, 1436-1438, 1984  
A:Title: Alpha-Cardiac actin is the major sarcomeric isoform expressed in embryonic avia  
A:Reference number: A43616; MUID:84223949; PMID:6729461  
A:Accession: C43616  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-86 <PAT>  
A:Cross-references: UNIPROT:Q90736; GB:K02173; NID:g211054; PIDN:AAA98513.1; PID:g211055  
C:Superfamily: actin  
C:Keywords: cytosol; methylated amino acid  
F:73/Modified site: 3'-methylhistidine (His) #status predicted

Query Match 100.0%; Score 80; DB 2; Length 86;  
Best Local Similarity 100.0%; Pred. No. 7.4e-07;  
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 FPSIVGRPRHOGVMV 15  
Db 31 FPSIVGRPRHOGVMV 45

### RESULT 2

actin 5C - fruit fly (Drosophila melanogaster)  
C:Species: Drosophila melanogaster  
C:Date: 30-Sep-1989 #sequence\_revision 30-Sep-1989 #text\_change 09-Jul-2004  
C:Accession: A28258  
R:Vigoreaux, J.O.; Tobin, S.L.  
Genes Dev. 1, 1161-1171, 1987  
A:Title: Stage-specific selection of alternative transcriptional initiation sites from t  
A:Reference number: A28258; MUID:88112795; PMID:3123314  
A:Accession: A28258  
A:Molecule type: mRNA  
A:Residues: 1-137 <VIG>  
A:Cross-references: UNIPROT:P10987  
A:Note: the authors translated the codon GAG for residue 96 as Gly  
C:Genetics:  
A:Gene: FlyBase:Act5C  
A:Cross-references: FlyBase:FBgn0000042  
C:Superfamily: actin  
C:Keywords: methylated amino acid  
F:74/Modified site: 3'-methylhistidine (His) #status predicted

Query Match 100.0%; Score 80; DB 2; Length 137;  
Best Local Similarity 100.0%; Pred. No. 1.2e-06;

```
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 FPSIVGRPRHQGVV 15
Db 32 FPSIVGRPRHQGVV 46

RESULT 3
I49465
A;Title: alpha-cardiac actin - mouse (fragment)
A;Species: Mus musculus (house mouse)
C;Date: 02-Jul-1996 #sequence_revision 02-Jul-1996 #text_change 09-Jul-2004
C;Accession: I49465; I49466
R;Garner, I.; Minty, A.J.; Alonso, S.; Barton, P.J.R.; Buckingham, M.E.
EMBO J. 5, 2559-2567, 1986
A;Title: A 5' duplication of the alpha-cardiac actin gene in BALB/c mice is associated w
A;Reference number: I49465; MUID:87053822; PMID:3023046
A;Accession: I49465
A;Status: translated from GB/EMBL/DBJ
A;Molecule type: DNA
A;Residues: 1-158 <RES>
A;Cross-references: UNIPROT:Q61273; GB:M26775; NID:g191646; PIDN:AAA37165.1; PID:g553858
A;Accession: I49466
A;Status: translated from GB/EMBL/DBJ
A;Molecule type: DNA
A;Residues: 8-158 <RES>
A;Cross-references: GB:M26776; NID:g191649; PIDN:AAA37166.1; PID:g553859
A;Experimental source: adult cardiac muscle, BALB/c mice
C;Genetics:
A;Introns: 50/3
C;Superfamily: actin
C;Keywords: cardiac muscle; heart

Query Match 100.0%; Score 80; DB 2; Length 158;
Best Local Similarity 100.0%; Pred. No. 1.4e-06;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 FPSIVGRPRHQGVV 15
Db 40 FPSIVGRPRHQGVV 54

RESULT 4
A61043
actin CA15 - sea squirt (Styela clava) (fragments)
C;Species: Styela clava
C;Date: 31-Dec-1993 #sequence_revision 31-Dec-1993 #text_change 09-Jul-2004
C;Accession: A61043
R;Beach, R.L.; Jeffery, W.R.
Dev. Genet. 11, 2-14, 1990
A;Title: Temporal and spatial expression of a cytoskeletal actin gene in the ascidian s
A;Reference number: A61043; MUID:90298580; PMID:2361333
A;Accession: A61043
A;Status: not compared with conceptual translation
A;Molecule type: mRNA
A;Residues: 1-213 <BEA>
A;Cross-references: UNIPROT:Q7M3Y7
C;Comment: This sequence is expressed in cells undergoing rapid cell division.
C;Superfamily: actin
C;Keywords: cytoskeleton; methylated amino acid; mitosis; structural protein
F;73/Modified site: 3'-methylhistidine (His) #status predicted

Query Match 100.0%; Score 80; DB 2; Length 213;
Best Local Similarity 100.0%; Pred. No. 1.9e-06;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 FPSIVGRPRHQGVV 15
Db 31 FPSIVGRPRHQGVV 45

RESULT 5
A03000
actin 3 - fruit fly (Drosophila melanogaster) (fragments)
C;Species: Drosophila melanogaster
C;Date: 31-Mar-1991 #sequence_revision 31-Mar-1991 #text_change 09-Jul-2004
C;Accession: A03000
R;Fyrberg, E.A.; Bond, B.J.; Hershey, N.D.; Mixter, K.S.; Davidson, N.
Cell 24, 107-116, 1981
A;Title: The actin genes of Drosophila: protein coding regions are highly conserved but :
A;Reference number: A03000; MUID:81210174; PMID:6263481
A;Accession: A03000
A;Molecule type: DNA
A;Residues: 1-308 <FYR>
A;Cross-references: UNIPROT:P02572
A;Note: there are 68 unsequenced residues between positions 160 and 161. Partial sequenc
A;Note: the authors translated the codon GTT for residue 263 as Ile
C;Genetics:
A;Gene: FlyBase:Act42A
A;Cross-references: FlyBase:FBgn0000043
A;Map position: 42A
C;Superfamily: actin
C;Keywords: methylated amino acid
F;74/Modified site: 3'-methylhistidine (His) #status predicted

Query Match 100.0%; Score 80; DB 2; Length 308;
Best Local Similarity 100.0%; Pred. No. 2.8e-06;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 FPSIVGRPRHQGVV 15
Db 32 FPSIVGRPRHQGVV 46

RESULT 6
S05430
actin beta - grass carp
C;Species: Ctenopharyngodon idella (grass carp)
C;Date: 07-Jun-1990 #sequence_revision 07-Jun-1990 #text_change 09-Jul-2004
C;Accession: S05430
R;Liu, Z.; Zhu, Z.; Roberg, K.; Faras, A.J.; Guise, K.S.; Kapuscinski, A.R.; Hackett, P.
Nucleic Acids Res. 17, 5850, 1989
A;Title: The beta-actin gene of carp (Ctenopharyngodon idella).
A;Reference number: S05430; MUID:89345185; PMID:2762162
A;Accession: S05430
A;Status: translation not shown
A;Molecule type: DNA
A;Residues: 1-328 <LIU>
A;Cross-references: UNIPROT:P83751; EMBL:M25013
C;Genetics:
A;Introns: 41/3; 121/3; 268/1
C;Superfamily: actin
C;Keywords: cytoskeleton; methylated amino acid
F;73/Modified site: 3'-methylhistidine (His) #status predicted

Query Match 100.0%; Score 80; DB 2; Length 328;
Best Local Similarity 100.0%; Pred. No. 3e-06;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 FPSIVGRPRHQGVV 15
Db 31 FPSIVGRPRHQGVV 45

RESULT 7
B25819
actin, fetal skeletal/adult cardiac muscle - mouse (fragment)
C;Species: Mus musculus (house mouse)
C;Date: 23-Aug-1987 #sequence_revision 23-Aug-1987 #text_change 09-Jul-2004
C;Accession: B25819
R;Alonso, S.; Minty, A.; Bourlet, Y.; Buckingham, M.
J. Mol. Evol. 23, 11-22, 1986
A;Title: Comparison of three actin-coding sequences in the mouse: evolutionary relations
A;Reference number: A25819; MUID:86200234; PMID:3084797
A;Accession: B25819
A;Molecule type: mRNA
```

```
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 FPSIVGRPRHQGVV 15
Db 32 FPSIVGRPRHQGVV 46

RESULT 3
I49465
A;Title: alpha-cardiac actin - mouse (fragment)
A;Species: Mus musculus (house mouse)
C;Date: 02-Jul-1996 #sequence_revision 02-Jul-1996 #text_change 09-Jul-2004
C;Accession: I49465; I49466
R;Garner, I.; Minty, A.J.; Alonso, S.; Barton, P.J.R.; Buckingham, M.E.
EMBO J. 5, 2559-2567, 1986
A;Title: A 5' duplication of the alpha-cardiac actin gene in BALB/c mice is associated w
A;Reference number: I49465; MUID:87053822; PMID:3023046
A;Accession: I49465
A;Status: translated from GB/EMBL/DBJ
A;Molecule type: DNA
A;Residues: 1-158 <RES>
A;Cross-references: UNIPROT:Q61273; GB:M26775; NID:g191646; PIDN:AAA37165.1; PID:g553858
A;Accession: I49466
A;Status: translated from GB/EMBL/DBJ
A;Molecule type: DNA
A;Residues: 8-158 <RES>
A;Cross-references: GB:M26776; NID:g191649; PIDN:AAA37166.1; PID:g553859
A;Experimental source: adult cardiac muscle, BALB/c mice
C;Genetics:
A;Introns: 50/3
C;Superfamily: actin
C;Keywords: cardiac muscle; heart

Query Match 100.0%; Score 80; DB 2; Length 158;
Best Local Similarity 100.0%; Pred. No. 1.4e-06;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 FPSIVGRPRHQGVV 15
Db 40 FPSIVGRPRHQGVV 54

RESULT 4
A61043
actin CA15 - sea squirt (Styela clava) (fragments)
C;Species: Styela clava
C;Date: 31-Dec-1993 #sequence_revision 31-Dec-1993 #text_change 09-Jul-2004
C;Accession: A61043
R;Beach, R.L.; Jeffery, W.R.
Dev. Genet. 11, 2-14, 1990
A;Title: Temporal and spatial expression of a cytoskeletal actin gene in the ascidian s
A;Reference number: A61043; MUID:90298580; PMID:2361333
A;Accession: A61043
A;Status: not compared with conceptual translation
A;Molecule type: mRNA
A;Residues: 1-213 <BEA>
A;Cross-references: UNIPROT:Q7M3Y7
C;Comment: This sequence is expressed in cells undergoing rapid cell division.
C;Superfamily: actin
C;Keywords: cytoskeleton; methylated amino acid; mitosis; structural protein
F;73/Modified site: 3'-methylhistidine (His) #status predicted

Query Match 100.0%; Score 80; DB 2; Length 213;
Best Local Similarity 100.0%; Pred. No. 1.9e-06;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 FPSIVGRPRHQGVV 15
Db 31 FPSIVGRPRHQGVV 45

RESULT 5
A03000
actin 3 - fruit fly (Drosophila melanogaster) (fragments)
C;Species: Drosophila melanogaster
C;Date: 31-Mar-1991 #sequence_revision 31-Mar-1991 #text_change 09-Jul-2004
C;Accession: A03000
R;Fyrberg, E.A.; Bond, B.J.; Hershey, N.D.; Mixter, K.S.; Davidson, N.
Cell 24, 107-116, 1981
A;Title: The actin genes of Drosophila: protein coding regions are highly conserved but :
A;Reference number: A03000; MUID:81210174; PMID:6263481
A;Accession: A03000
A;Molecule type: DNA
A;Residues: 1-308 <FYR>
A;Cross-references: UNIPROT:P02572
A;Note: there are 68 unsequenced residues between positions 160 and 161. Partial sequenc
A;Note: the authors translated the codon GTT for residue 263 as Ile
C;Genetics:
A;Gene: FlyBase:Act42A
A;Cross-references: FlyBase:FBgn0000043
A;Map position: 42A
C;Superfamily: actin
C;Keywords: methylated amino acid
F;74/Modified site: 3'-methylhistidine (His) #status predicted

Query Match 100.0%; Score 80; DB 2; Length 308;
Best Local Similarity 100.0%; Pred. No. 2.8e-06;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 FPSIVGRPRHQGVV 15
Db 32 FPSIVGRPRHQGVV 46

RESULT 6
S05430
actin beta - grass carp
C;Species: Ctenopharyngodon idella (grass carp)
C;Date: 07-Jun-1990 #sequence_revision 07-Jun-1990 #text_change 09-Jul-2004
C;Accession: S05430
R;Liu, Z.; Zhu, Z.; Roberg, K.; Faras, A.J.; Guise, K.S.; Kapuscinski, A.R.; Hackett, P.
Nucleic Acids Res. 17, 5850, 1989
A;Title: The beta-actin gene of carp (Ctenopharyngodon idella).
A;Reference number: S05430; MUID:89345185; PMID:2762162
A;Accession: S05430
A;Status: translation not shown
A;Molecule type: DNA
A;Residues: 1-328 <LIU>
A;Cross-references: UNIPROT:P83751; EMBL:M25013
C;Genetics:
A;Introns: 41/3; 121/3; 268/1
C;Superfamily: actin
C;Keywords: cytoskeleton; methylated amino acid
F;73/Modified site: 3'-methylhistidine (His) #status predicted

Query Match 100.0%; Score 80; DB 2; Length 328;
Best Local Similarity 100.0%; Pred. No. 3e-06;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 FPSIVGRPRHQGVV 15
Db 31 FPSIVGRPRHQGVV 45

RESULT 7
B25819
actin, fetal skeletal/adult cardiac muscle - mouse (fragment)
C;Species: Mus musculus (house mouse)
C;Date: 23-Aug-1987 #sequence_revision 23-Aug-1987 #text_change 09-Jul-2004
C;Accession: B25819
R;Alonso, S.; Minty, A.; Bourlet, Y.; Buckingham, M.
J. Mol. Evol. 23, 11-22, 1986
A;Title: Comparison of three actin-coding sequences in the mouse: evolutionary relations
A;Reference number: A25819; MUID:86200234; PMID:3084797
A;Accession: B25819
A;Molecule type: mRNA
```



A:Residues: 1-349 <ALO>  
A:Cross-references: UNIPROT:Q61275; GB:X03767; GB:J00381; GB:M10652; NID:G49869; PIDN:CA  
C:Superfamily: actin  
C:Keywords: cardiac muscle; heart; methylated amino acid; muscle; skeletal muscle  
F:47/Modified site: 3'-methylhistidine (His) #status predicted

Query Match 100.0%; Score 80; DB 2; Length 349;  
Best Local Similarity 100.0%; Pred. No. 3.2e-06;  
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 FPSIVGRPRHQGMV 15  
|||||  
Db 5 FPSIVGRPRHQGMV 19  
|||||

RESULT 8  
A26559  
actin type 5, cytosolic - chicken  
C:Species: Gallus gallus (chicken)  
C>Date: 09-Sep-1987 #sequence\_revision 09-Sep-1987 #text\_change 05-Dec-1997  
C:Accession: A26559  
R:Bergsma, D.J.; Chang, K.S.; Schwartz, R.J.  
Mol. Cell. Biol. 5, 1151-1162, 1985  
A:Reference number: A26559; MUID:85213487; PMID:4000121  
A:Accession: A26559  
A:Molecule type: DNA  
A:Residues: 1-362 <BER>  
C:Superfamily: actin  
C:Keywords: cytosol; methylated amino acid  
F:74/Modified site: 3'-methylhistidine (His) #status predicted

Query Match 100.0%; Score 80; DB 2; Length 362;  
Best Local Similarity 100.0%; Pred. No. 3.4e-06;  
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 FPSIVGRPRHQGMV 15  
|||||  
Db 32 FPSIVGRPRHQGMV 46  
|||||

RESULT 9  
ATBOG  
actin beta - bovine (tentative sequence)  
C:Species: Bos primigenius taurus (cattle)  
C>Date: 30-Sep-1991 #sequence\_revision 30-Sep-1991 #text\_change 09-Jul-2004  
C:Accession: E14185; A39105; A02999; A14185  
R:Vandekerckhove, J.; Weber, K.  
Eur. J. Biochem. 90, 451-462, 1978  
A:Title: Actin amino-acid sequences. Comparison of actins from calf thymus, bovine brain  
A:Reference number: A14185; MUID:79045349; PMID:213279  
A:Accession: E14185  
A:Molecule type: protein  
A:Residues: 1-374 <VAN>  
A:Cross-references: UNIPROT:P60712  
A:Note: only peptides that differed in composition from the corresponding peptides of ra  
R:Degen, J.L.; Neubauer, M.G.; Degen, S.J.P.; Seyfried, C.E.; Morris, D.R.  
J. Biol. Chem. 258, 12153-12162, 1983  
A:Title: Regulation of protein synthesis in mitogen-activated bovine lymphocytes. Analysis  
A:Reference number: A39105; MUID:84032385; PMID:6195151  
A:Accession: A39105  
A:Molecule type: mRNA  
A:Residues: 76-227;344-374 <DEG>  
A:Cross-references: GB:K00622; GB:K00623  
A:Note: actins beta and gamma were not distinguished in this study  
C:Comment: Vertebrate nonmuscle cells contain two highly conserved cytoskeletal actins,  
C:Superfamily: actin  
C:Keywords: blocked amino end; cell motility; cytoskeleton; methylated amino acid; micro  
F:1/Modified site: blocked amino end (Asp) (probably acetylated) #status experimental  
F:72/Modified site: 3'-methylhistidine (His) #status predicted

Query Match 100.0%; Score 80; DB 1; Length 374;  
Best Local Similarity 100.0%; Pred. No. 3.5e-06;  
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 FPSIVGRPRHQGMV 15  
|||||  
Db 30 FPSIVGRPRHQGMV 44  
|||||

RESULT 10  
ATBOG  
actin gamma - bovine (tentative sequence)  
C:Species: Bos primigenius taurus (cattle)  
C>Date: 30-Sep-1991 #sequence\_revision 30-Sep-1991 #text\_change 09-Jul-2004  
C:Accession: B14185; A02999  
R:Vandekerckhove, J.; Weber, K.  
Eur. J. Biochem. 90, 451-462, 1978  
A:Title: Actin amino-acid sequences. Comparison of actins from calf thymus, bovine brain  
A:Reference number: A14185; MUID:79045349; PMID:213279  
A:Accession: B14185  
A:Molecule type: protein  
A:Residues: 1-374 <VAN>  
A:Cross-references: UNIPROT:P02571  
A:Note: only peptides that differed in composition from the corresponding peptides of ra  
C:Comment: Vertebrate nonmuscle cells contain two highly conserved cytoskeletal actins,  
C:Superfamily: actin  
C:Keywords: blocked amino end; cell motility; cytoskeleton; methylated amino acid; micro  
F:1/Modified site: blocked amino end (Glu) (probably acetylated) #status experimental  
F:72/Modified site: 3'-methylhistidine (His) #status predicted

Query Match 100.0%; Score 80; DB 1; Length 374;  
Best Local Similarity 100.0%; Pred. No. 3.5e-06;  
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 FPSIVGRPRHQGMV 15  
|||||  
Db 30 FPSIVGRPRHQGMV 44  
|||||

RESULT 11  
JC5818  
gamma-actin - human  
C:Species: Homo sapiens (man)  
C>Date: 20-Feb-1998 #sequence\_revision 13-Mar-1998 #text\_change 09-Jul-2004  
C:Accession: JC5818; PC4501  
R:Hauschildt, S.; Schwarz, C.; Heine, H.; Ulmer, A.J.; Flad, H.D.; Rietschel, E.T.; Jens  
Biochem. Biophys. Res. Commun. 241, 670-674, 1997  
A:Title: Actin: A target of lipopolysaccharide-induced phosphorylation in human monocytes  
A:Reference number: JC5818; MUID:98096379; PMID:9434766  
A:Accession: JC5818  
A:Molecule type: protein  
A:Residues: 1-374 <HAU>  
A:Cross-references: UNIPROT:P02571  
A:Experimental source: monocyte  
A:Accession: PC4501  
A:Molecule type: protein  
A:Residues: 1-61;84-112;147-190;196-209;215-253;335-358 <HA2>  
A:Experimental source: monocyte  
C:Comment: This protein is involved in a signal transduction that eventually leads to mo  
C:Superfamily: actin

Query Match 100.0%; Score 80; DB 2; Length 374;  
Best Local Similarity 100.0%; Pred. No. 3.5e-06;  
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 FPSIVGRPRHQGMV 15  
|||||  
Db 30 FPSIVGRPRHQGMV 44  
|||||

RESULT 12  
ATBOG  
actin, aortic smooth muscle - bovine  
C:Species: Bos primigenius taurus (cattle)  
C>Date: 30-Jun-1987 #sequence\_revision 30-Jun-1987 #text\_change 06-Sep-1996  
C:Accession: A02997; S13480

R;Vandekerckhove, J.; Weber, K.  
Differentiation 14, 123-133, 1979  
A;Title: The complete amino acid sequence of actins from bovine aorta, bovine heart, bovine thymus.  
A;Reference number: A02997; MUID:80047657; PMID:499690  
A;Accession: A02997  
A;Molecule type: protein  
A;Residues: 1-375 <VAN>  
R;Zevgoliss, V.G.; Sotiropoulos, T.G.; Evangelopoulos, A.E.  
Biochim. Biophys. Acta 1091, 222-230, 1991  
A;Title: Phosphorylase kinase from bovine stomach smooth muscle: a Ca(2+)-dependent protein.  
A;Reference number: S13480; MUID:91137633; PMID:1995080  
A;Accession: S13480  
A;Molecule type: protein  
A;Residues: 40-49 <ZEV>  
A;Experimental source: stomach  
A;Note: This material appears to be actin of aortic smooth muscle type or a related molecule.  
C;Superfamily: actin  
C;Keywords: acetylated amino end; methylated amino acid; muscle contraction  
F;1/Modified site: acetylated amino end (Glu) #status predicted  
F;73/Modified site: 3'-methylhistidine (His) #status experimental

Query Match 100.0%; Score 80; DB 1; Length 375;  
Best Local Similarity 100.0%; Pred. No. 3.5e-06;  
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 FPSIVGRPRHQGMV 15  
Db 31 FPSIVGRPRHQGMV 45

RESULT 13

ATRB  
actin, skeletal muscle - rabbit  
N;Alternate names: F-actin  
C;Species: Oryctolagus cuniculus (domestic rabbit)  
C;Date: 30-Sep-1991 #sequence revision 30-Sep-1991 #text change 07-May-1999  
C;Accession: A92182; A14185; S65873; S70610; S44393; A02994  
R;Collins, J.H.; Elzinga, M.  
J. Biol. Chem. 250, 5915-5920, 1975  
A;Title: The primary structure of actin from rabbit skeletal muscle. Completion and analysis.  
A;Reference number: A92182; MUID:75211334; PMID:1150665  
A;Accession: A92182  
A;Molecule type: protein  
A;Residues: 1-2, 1, 4, D, 6-11, D, 13-73, W, 74-78, 80-234, 236-308, T, 310-375 <COL>  
A;Note: This is the final paper in a series  
A;Note: This sequence has been revised in references A14185 and A90406  
R;Vandekerckhove, J.; Weber, K.  
Eur. J. Biochem. 90, 451-462, 1978  
A;Title: Actin amino-acid sequences. Comparison of actins from calf thymus, bovine brain and rabbit skeletal muscle.  
A;Reference number: A14185; MUID:79045349; PMID:213279  
A;Accession: A14185  
A;Molecule type: protein  
A;Residues: 1-18; 69-84 <VAN>  
R;Lu, R.C.; Elzinga, M.  
Biochemistry 16, 5801-5806, 1977  
A;Title: Partial amino acid sequence of brain actin and its homology with muscle actin.  
A;Reference number: A90406; MUID:78060866; PMID:588555  
A;Contents: annotation  
A;Note: residue 235 has been added and residue 309 has been revised  
R;Vandat, A.; Miller, C.; Phillips, M.; Muhltad, A.; Reisler, E.  
FEBS Lett. 365, 149-151, 1995  
A;Title: A novel 27/16 kDa form of subunit cleaved actin: structural and functional characterization.  
A;Reference number: S65873; MUID:95300963; PMID:7781768  
A;Accession: S65873  
A;Molecule type: protein  
A;Residues: 235-241 <VAN>  
R;Strzelecka-Golaszewska, H.; Wozniak, A.; Hult, T.; Lindberg, U.  
Biochem. J. 316, 713-721, 1996  
A;Title: Effects of the type of divalent cation, Ca2+ or Mg2+, bound at the high-affinity site on the actin monomer.  
A;Reference number: S70610; MUID:96265033; PMID:8670143  
A;Accession: S70610  
A;Molecule type: protein

A;Residues: 48-54; 68-72; 235-243 <STR>  
A;Experimental source: skeletal muscle  
R;Bertrand, R.; Derancourt, J.; Kassab, R.  
FEBS Lett. 345, 113-119, 1994  
A;Title: The covalent maleimidebenzoyl-actin-myosin head complex. Cross-linking of the 5' and 10' regions of the actin monomer.  
A;Reference number: S44393; MUID:94259162; PMID:8200441  
A;Accession: S44393  
A;Molecule type: protein  
A;Residues: 48-64 <BER>  
A;Experimental source: skeletal muscle  
C;Superfamily: actin  
C;Keywords: acetylated amino end; ATP binding; methylated amino acid; muscle contraction  
F;1-375/Product: actin #status experimental <MAT>  
F;1/Modified site: acetylated amino end (Asp) #status experimental  
F;73/Modified site: 3'-methylhistidine (His) #status experimental

Query Match 100.0%; Score 80; DB 1; Length 375;  
Best Local Similarity 100.0%; Pred. No. 3.5e-06;  
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 FPSIVGRPRHQGMV 15  
Db 31 FPSIVGRPRHQGMV 45

RESULT 14

ATRIC  
actin beta - rat  
C;Species: Rattus norvegicus (Norway rat)  
C;Date: 13-Jun-1983 #sequence revision 13-Jun-1983 #text change 22-Jun-1999  
C;Accession: A38571; A02999  
R;Nudel, U.; Zakut, R.; Shani, M.; Neuman, S.; Levy, Z.; Yaffe, D.  
Nucleic Acids Res. 11, 1759-1771, 1983  
A;Title: The nucleotide sequence of the rat cytoplasmic beta-actin gene.  
A;Reference number: A38571; MUID:83168920; PMID:6300777  
A;Accession: A38571  
A;Molecule type: DNA  
A;Residues: 1-375 <NUD>  
A;Cross-references: GB:J00691; NID:q202653; PIDN:AAA40657.1; PID:q202654  
C;Comment: Vertebrate nonmuscle cells contain two highly conserved cytoskeletal actins, actin beta and actin gamma.  
C;Genetics: 41/3; 121/3; 268/1; 328/3  
C;Superfamily: actin  
C;Keywords: cell motility; cytoskeleton; methylated amino acid; microfilament; mitosis; F-actin  
F;2-375/Product: actin beta #status predicted <MAT>  
F;73/Modified site: 3'-methylhistidine (His) #status predicted

Query Match 100.0%; Score 80; DB 1; Length 375;  
Best Local Similarity 100.0%; Pred. No. 3.5e-06;  
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 FPSIVGRPRHQGMV 15  
Db 31 FPSIVGRPRHQGMV 45

RESULT 15

A48324  
actin beta, cytoskeletal - common carp  
C;Species: Cyprinus carpio (common carp)  
C;Date: 03-Feb-1994 #sequence revision 11-Apr-1997 #text change 09-Jul-2004  
C;Accession: A48324  
R;Liu, Z.; Zhu, Z.; Roberg, K.; Faras, A.; Guise, K.; Kapuscinski, A.R.; Hackett, P.B.  
DNA Seq. 1, 125-136, 1990  
A;Title: Isolation and characterization of beta-actin gene of carp (Cyprinus carpio).  
A;Reference number: A48324; MUID:92190540; PMID:2134183  
A;Accession: A48324  
A;Molecule type: DNA  
A;Residues: 1-375 <LIU>  
A;Cross-references: UNIPROT:P83750; GB:M24113; NID:g213041; PIDN:AAA68886.1; PID:g213042  
A;Note: The authors translated the codon TTC for residue 21 as Pro, AAG for residue 50 as Pro  
A;Note: the authors failed to translated the codon GGT for residue 42 as Gly

C;Comment: Vertebrate nonmuscle cells contain two highly conserved cytoskeletal actins,  
C;Genetics:  
A;Introns: 41/3; 121/3; 268/3; 328/3  
C;Superfamily: actin  
C;Keywords: acetylated amino end; cell motility; cytoskeleton; methylated amino acid; mi  
F;2-375/Product: actin beta, cytoskeletal #status predicted <MAT>-  
F;2/Modified site: acetylated amino end (Asp) (in mature form) #status predicted  
F;73/Modified site: 3'-methylhistidine (His) #status predicted

Query Match 100.0%; Score 80; DB 1; Length 375;  
Best Local Similarity 100.0%; Pred. No. 3.5e-06;  
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 FPSIVGRPRHOGVMV 15  
|||||  
Db 31 FPSIVGRPRHOGVMV 45  
|||||

Search completed: April 8, 2005, 10:53:04  
Job time : 9.33333 secs

**THIS PAGE BLANK (USPTO)**

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: April 8, 2005, 09:58:56 ; Search time 57.8667 Seconds  
(without alignments)  
132.739 Million cell updates/sec

Title: US-09-423-351C-2  
Perfect score: 80  
Sequence: 1 FPSIVGRPRHQGVNV 15

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 1612378 seqs, 512079187 residues

Total number of hits satisfying chosen parameters: 1612378

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : UniProt\_03 : \*  
1: uniprot\_sprot : \*  
2: uniprot\_trembl : \*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	80	100.0	26	Q6Q298	Q6Q298 ovis aries
2	80	100.0	53	Q6AW41	Q6AW41 bombyx mori
3	80	100.0	78	Q61375	Q61375 schmidtea p
4	80	100.0	78	Q61376	Q61376 schmidtea p
5	80	100.0	78	Q61377	Q61377 schmidtea p
6	80	100.0	79	Q95L83	Q95L83 oryctolagus
7	80	100.0	86	Q90736	Q90736 gallus gall
8	80	100.0	91	Q72211	Q72211 salvelinus
9	80	100.0	96	Q86CV4	Q86cv4 drosophila
10	80	100.0	96	Q86FV7	Q86fv7 drosophila
11	80	100.0	99	Q9QZB4	Q9Qzb4 cavia porce
12	80	100.0	104	Q86SB7	Q86sb7 gryllus bim
13	80	100.0	108	Q18548	Q18548 lytechinus
14	80	100.0	108	Q18550	Q18550 lytechinus
15	80	100.0	121	Q83G71	Q83g71 oreochromis
16	80	100.0	121	Q78CQ1	Q78cq1 pseudopleur
17	80	100.0	129	Q6VBE4	Q6vbe4 locusta mig
18	80	100.0	130	Q93298	Q93298 pseudopleur
19	80	100.0	139	Q9DFK2	Q9dfk2 gillichthys
20	80	100.0	144	Q90YX9	Q90yx9 fundulus he
21	80	100.0	150	Q804Y9	Q804y9 ictalurus p
22	80	100.0	151	Q61274	Q61274 mus musculu
23	80	100.0	151	Q9PSV5	Q9psv5 oryzias lat
24	80	100.0	158	Q61273	Q61273 mus musculu
25	80	100.0	164	1 ACT SPOLI	Q11212 spodoptera
26	80	100.0	165	Q17477	Q17477 haliotis ru
27	80	100.0	184	Q11209	Q11209 canis fami
28	80	100.0	186	Q8X190	Q8x190 paxillus in
29	80	100.0	189	Q8SPX4	Q8spx4 canis fami
30	80	100.0	191	Q64G12	Q64g12 oxyuranus s
31	80	100.0	198	Q6Y250	Q6y250 pagrus majo

32	80	100.0	202	Q810R5	Q810r5 mus musculu
33	80	100.0	212	Q6ZYL2	Q6zyl2 arion lusit
34	80	100.0	213	Q7M3V7	Q7m3v7 styela clav
35	80	100.0	225	Q64G13	Q64g13 oxyuranus s
36	80	100.0	235	Q7XB23	Q7xb23 lotharella
37	80	100.0	235	Q7XB24	Q7xb24 lotharella
38	80	100.0	235	Q7XB25	Q7xb25 lotharella
39	80	100.0	236	Q6RXK3	Q6rxk3 penaeus van
40	80	100.0	237	Q7XB21	Q7xb21 lotharella
41	80	100.0	237	Q7XB22	Q7xb22 lotharella
42	80	100.0	238	Q7XB20	Q7xb20 lotharella
43	80	100.0	239	Q7XB28	Q7xb28 chlorarachn
44	80	100.0	254	Q7Z7J6	Q7z7j6 homo sapien
45	80	100.0	273	Q7Q7K4	Q7q7k4 anopheles g

## ALIGNMENTS

### RESULT 1

Q6Q298 PRELIMINARY; PRT; 26 AA.  
AC Q6Q298;  
DT 05-JUL-2004 (TrEMBLrel. 27, Created)  
DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)  
DT 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)  
DE Beta actin (Fragment).  
OS Ovis aries (Sheep).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;  
OC Caprinae; Ovis.  
OX NCBI\_TaxID=9940;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC TISSUE=Abomasum;  
RA Shanmuganathan T., Hickford J.G.H., Savill M., Sykes A.R.;  
RL Submitted (MAR-2004) to the EMBL/GenBank/DBJ databases.  
FR EMBL; AY566300; AAS68014.1; -.  
FT NON\_TER 1 1  
FT NON\_TER 26 26  
SQ SEQUENCE 26 AA; 2802 MW; C75DA20C92E36C4B CRC64;

Query Match 100.0%; Score 80; DB 2; Length 26;  
Best Local Similarity 100.0%; Pred. No. 8.2e-07;  
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 FPSIVGRPRHQGVNV 15  
Db 2 FPSIVGRPRHQGVNV 16  
|||||  
|||||

### RESULT 2

Q6AW41 PRELIMINARY; PRT; 53 AA.  
AC Q6AW41;  
DT 25-OCT-2004 (TrEMBLrel. 28, Created)  
DT 25-OCT-2004 (TrEMBLrel. 28, Last sequence update)  
DT 25-OCT-2004 (TrEMBLrel. 28, Last annotation update)  
DE Cytoplasmic actin (Fragment).  
OS Bombyx mori (Silk moth).  
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;  
OC Neoptera; Endopterygota; Lepidoptera; Glossata; Ditryaia; Bombycoidea;  
OC Bombycidae; Bombyx.  
OX NCBI\_TaxID=7091;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=Daizo; TISSUE=Silk gland;  
RA Shioimi K., Kajiuira Z., Nakagaki M., Yamashita O.;  
RL "Baculovirus-mediated efficient gene transfer into the central nervous system of the silkworm, Bombyx mori.";  
RL Nihon Sanehigaku Zaeshi 72:149-155 (2003).  
RN [2]  
RP SEQUENCE FROM N.A.

```

RC STRAIN=Daizo; TISSUE=Silk gland;
RA Shomi K.;
RL Submitted (AUG-2004) to the EMBL/GenBank/DBJ databases.
RM EMBL; AB186491; BAD35130.1; -.
DR InterPro; IPR004000; Actin_like.
DR Pfam; PF00022; Actin; 1.
FT NON TER 53
SQ SEQUENCE 53 AA; 5465 MW; 227AA8B4872EBA86 CRC64;

Query Match 100.0%; Score 80; DB 2; Length 53;
Best Local Similarity 100.0%; Pred. No. 1.8e-06;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 FPSIVGRPRHQGVWV 15
Db 32 FPSIVGRPRHQGVWV 46

RESULT 3
O61375 PRELIMINARY; PRT; 78 AA.
ID O61375
AC O61375;
DT 01-AUG-1998 (TrEMBLrel. 07, Created)
DT 01-AUG-1998 (TrEMBLrel. 07, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE Actin 1 (Fragment).
GN Name=DpAct1;
OS Schmidtea polychroa.
OC Eukaryota; Metazoa; Platyhelminthes; Turbellaria; Seriata; Tricladida;
OC Paludicola; Dugesiiidae; Schmidtea.
OX NCBI_TaxID=50054;
RN [1]
RP SEQUENCE FROM N.A.
MEDLINE=98186847; PubMed=9518500; DOI=10.1093/nar/26.8.2031;
RA Fagotti A., Gabbiani G., Pascolini R., Neuville P.;
RT "Multiple isoform recovery (MIR)-PCR: a simple method for the
isolation of related mRNA isoforms.";
RL Nucleic Acids Res. 26:2031-2033(1998).
CC -!- FUNCTION: Actins are highly conserved proteins that are involved
in various types of cell motility and are ubiquitously expressed
in all eukaryotic cells (By similarity).
CC -!- SIMILARITY: Belongs to the actin family.
DR EMBL; AF027161; AAC38981.1; -.
DR HSSP; P02577; INMI.
DR GO; GO:0005884; C:actin filament; IEA.
DR GO; GO:0003774; F:motor activity; IEA.
DR GO; GO:0005200; F:structural constituent of cytoskeleton; IEA.
DR InterPro; IPR004001; Actin_like.
DR Pfam; PF00022; Actin; 1.
DR PRINTS; PR00190; ACTIN.
DR PROSITE; PS00406; ACTINS_1; 1.
KW Structural protein.
FT NON TER 78
SQ SEQUENCE 78 AA; 8205 MW; 572BAB2698C24404 CRC64;

Query Match 100.0%; Score 80; DB 2; Length 78;
Best Local Similarity 100.0%; Pred. No. 2.7e-06;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 FPSIVGRPRHQGVWV 15
Db 32 FPSIVGRPRHQGVWV 46

RESULT 4
O61376 PRELIMINARY; PRT; 78 AA.
ID O61376
AC O61376;
DT 01-AUG-1998 (TrEMBLrel. 07, Created)
DT 01-AUG-1998 (TrEMBLrel. 07, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE Actin 2 (Fragment).

```

```

GN Name=DpAct2;
OS Schmidtea polychroa.
OC Eukaryota; Metazoa; Platyhelminthes; Turbellaria; Seriata; Tricladida;
OC Paludicola; Dugesiiidae; Schmidtea.
OX NCBI_TaxID=50054;
RN [1]
RP SEQUENCE FROM N.A.
MEDLINE=98186847; PubMed=9518500; DOI=10.1093/nar/26.8.2031;
RA Fagotti A., Gabbiani G., Pascolini R., Neuville P.;
RT "Multiple isoform recovery (MIR)-PCR: a simple method for the
isolation of related mRNA isoforms.";
RL Nucleic Acids Res. 26:2031-2033(1998).
CC -!- FUNCTION: Actins are highly conserved proteins that are involved
in various types of cell motility and are ubiquitously expressed
in all eukaryotic cells (By similarity).
CC -!- SIMILARITY: Belongs to the actin family.
DR EMBL; AF027162; AAC38982.1; -.
DR HSSP; P02577; INMI.
DR GO; GO:0005884; C:actin filament; IEA.
DR GO; GO:0003774; F:motor activity; IEA.
DR GO; GO:0005200; F:structural constituent of cytoskeleton; IEA.
DR InterPro; IPR004001; Actin_like.
DR Pfam; PF00022; Actin; 1.
DR PRINTS; PR00190; ACTIN.
DR PROSITE; PS00406; ACTINS_1; 1.
KW Structural protein.
FT NON TER 78
SQ SEQUENCE 78 AA; 8255 MW; B9DFF9B28E4ADB0E CRC64;

Query Match 100.0%; Score 80; DB 2; Length 78;
Best Local Similarity 100.0%; Pred. No. 2.7e-06;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 FPSIVGRPRHQGVWV 15
Db 32 FPSIVGRPRHQGVWV 46

RESULT 5
O61377 PRELIMINARY; PRT; 78 AA.
ID O61377
AC O61377;
DT 01-AUG-1998 (TrEMBLrel. 07, Created)
DT 01-AUG-1998 (TrEMBLrel. 07, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE Actin 3 (Fragment).
GN Name=DpAct3;
OS Schmidtea polychroa.
OC Eukaryota; Metazoa; Platyhelminthes; Turbellaria; Seriata; Tricladida;
OC Paludicola; Dugesiiidae; Schmidtea.
OX NCBI_TaxID=50054;
RN [1]
RP SEQUENCE FROM N.A.
MEDLINE=98186847; PubMed=9518500; DOI=10.1093/nar/26.8.2031;
RA Fagotti A., Gabbiani G., Pascolini R., Neuville P.;
RT "Multiple isoform recovery (MIR)-PCR: a simple method for the
isolation of related mRNA isoforms.";
RL Nucleic Acids Res. 26:2031-2033(1998).
CC -!- FUNCTION: Actins are highly conserved proteins that are involved
in various types of cell motility and are ubiquitously expressed
in all eukaryotic cells (By similarity).
CC -!- SIMILARITY: Belongs to the actin family.
DR EMBL; AF027163; AAC38983.1; -.
DR HSSP; P02577; INMI.
DR GO; GO:0005884; C:actin filament; IEA.
DR GO; GO:0003774; F:motor activity; IEA.
DR GO; GO:0005200; F:structural constituent of cytoskeleton; IEA.
DR InterPro; IPR004001; Actin.
DR Pfam; PF00022; Actin; 1.
DR PRINTS; PR00190; ACTIN.
DR PROSITE; PS00406; ACTINS_1; 1.

```

```

KW Structural protein. 78
FT NON TER 78
SQ SEQUENCE 78 AA; 8299 MW; B9DF1108E4ADBOE CRC64;

Query Match 100.0%; Score 80; DB 2; Length 78;
Best Local Similarity 100.0%; Pred. No. 2.7e-06;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 FPSIVGRPRHQGVV 15
Db 32 FPSIVGRPRHQGVV 46

RESULT 6
Q95L83
ID Q95L83 PRELIMINARY; PRT; 79 AA.
AC Q95L83;
DT 01-DEC-2001 (TrEMBLrel. 19, Created)
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE Beta-actin (Fragment).
OS Oryctolagus cuniculus (Rabbit).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Lagomorpha; Leporidae; Oryctolagus.
OX NCBI_TaxID=9986;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=23306422; PubMed=12388084;
RA Wang D.-A., Du H., Jaggar J.H., Brindley D.N., Tigyi G.J.,
RT "Injury-elicited differential transcriptional regulation of
RT phospholipid growth factor receptors in the cornea.";
RL Am. J. Physiol. Cell Physiol. 283:C1646-C1654(2002).
CC -!- FUNCTION: Actins are highly conserved proteins that are involved
CC in various types of cell motility and are ubiquitously expressed
CC in all eukaryotic cells (By similarity).
CC -!- SIMILARITY: Belongs to the actin family.
DR EMBL; AF04278; AAL01885.1; -.
DR HSSP; P02577; 1NM1.
DR GO; GO:0005884; C:actin filament; IEA.
DR GO; GO:0003774; F:motor activity; IEA.
DR GO; GO:0005200; F:structural constituent of cytoskeleton; IEA.
DR InterPro; IPR004001; Actin.
DR Pfam; PF00022; Actin; 1.
DR PRINTS; PR00190; ACTIN.
DR PROSITE; PS00406; ACTINS_1; 1.
DR Structural protein. 1
FT NON TER 1
FT NON TER 79
SQ SEQUENCE 79 AA; 8420 MW; 05056E4E2AC092F8 CRC64;

Query Match 100.0%; Score 80; DB 2; Length 79;
Best Local Similarity 100.0%; Pred. No. 2.7e-06;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 FPSIVGRPRHQGVV 15
Db 33 FPSIVGRPRHQGVV 47

RESULT 7
Q90736
ID Q90736 PRELIMINARY; PRT; 86 AA.
AC Q90736;
DT 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE Beta-actin (Fragment).
OS Gallus gallus (Chicken).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Archosauria; Aves; Neognathae; Galliformes; Phasianinae;
OC Gallus.

```

```

OX NCBI_TaxID=9031;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=84223949; PubMed=6729461;
RA Paterson B.M., Eldridge J.D.;
RT "alpha-Cardiac actin is the major sarcomeric isoform expressed in
RT embryonic avian skeletal muscle.";
RL Science 224:1436-1438(1984).
CC -!- FUNCTION: Actins are highly conserved proteins that are involved
CC in various types of cell motility and are ubiquitously expressed
CC in all eukaryotic cells (By similarity).
CC -!- SIMILARITY: Belongs to the actin family.
DR EMBL; K02173; AAA98513.1; -.
DR PIR; C43616; C43616.
DR HSSP; P02577; 1NM1.
DR GO; GO:0005884; C:actin filament; IEA.
DR GO; GO:0003774; F:motor activity; IEA.
DR GO; GO:0005200; F:structural constituent of cytoskeleton; IEA.
DR Pfam; PF00022; Actin; 1.
DR PRINTS; PR00190; ACTIN.
DR PROSITE; PS00406; ACTINS_1; 1.
DR Structural protein. 86
FT NON TER 86
FT NON TER 91
SQ SEQUENCE 86 AA; 9351 MW; A55285196A328B6E CRC64;

Query Match 100.0%; Score 80; DB 2; Length 86;
Best Local Similarity 100.0%; Pred. No. 2.9e-06;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 FPSIVGRPRHQGVV 15
Db 31 FPSIVGRPRHQGVV 45

RESULT 8
Q7Z211
ID Q7Z211 PRELIMINARY; PRT; 91 AA.
AC Q7Z211;
DT 01-JUN-2003 (TrEMBLrel. 24, Created)
DT 01-JUN-2003 (TrEMBLrel. 24, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Beta-actin (Fragment).
OS Salvelinus alpinus (Arctic char).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Euteleostei;
OC Protacanthopterygii; Salmoniformes; Salmonidae; Salvelinus.
OX NCBI_TaxID=8036;
RN [1]
RP SEQUENCE FROM N.A.
RA McGowan C., Davidson E.A., Davidson W.S.;
RL Submitted (MAR-2003) to the EMBL/GenBank/DBJ databases.
CC -!- FUNCTION: Actins are highly conserved proteins that are involved
CC in various types of cell motility and are ubiquitously expressed
CC in all eukaryotic cells (By similarity).
CC -!- SIMILARITY: Belongs to the actin family.
DR EMBL; AY262761; AAP31127.1; -.
DR HSSP; P02577; 1NM1.
DR GO; GO:0005884; C:actin filament; IEA.
DR GO; GO:0003774; F:motor activity; IEA.
DR GO; GO:0005200; F:structural constituent of cytoskeleton; IEA.
DR InterPro; IPR004001; Actin.
DR InterPro; IPR004000; Actin_like.
DR Pfam; PF00022; Actin; 1.
DR PRINTS; PR00190; ACTIN.
DR PROSITE; PS00406; ACTINS_1; 1.
DR Structural protein. 1
FT NON TER 1
FT NON TER 91
SQ SEQUENCE 91 AA; 10063 MW; 7118DB6663CD895C CRC64;

Query Match 100.0%; Score 80; DB 2; Length 91;
Best Local Similarity 100.0%; Pred. No. 3.1e-06;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```
Qy 1 FPSIVGRPRHQVMV 15
Db 29 FPSIVGRPRHQVMV 43

RESULT 9
Q86CV4 PRELIMINARY; PRT; 96 AA.
ID Q86CV4
AC Q86CV4;
DT 01-JUN-2003 (TrEMBLrel. 24, Created)
DT 01-JUN-2003 (TrEMBLrel. 24, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Actin E2 (Fragment)
GN Name=ActE2;
OS Drosophila novamexicana.
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
OX NCBI_TaxID=47314;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=15010-1031.0;
EX MEDLINE=22480340; PubMed=12574518; DOI=10.1073/pnas.0336368100;
RA Wittkopp P.J., Williams B.L., Selegue J.E., Carroll S.B.;
RT "Drosophila pigmentation evolution: divergent genotypes underlying
convergent phenotypes.";
RL Proc. Natl. Acad. Sci. U.S.A. 100:1808-1813(2003).
CC -!- FUNCTION: Actins are highly conserved proteins that are involved
in various types of cell motility and are ubiquitously expressed
in all eukaryotic cells (By similarity).
CC -!- SIMILARITY: Belongs to the actin family.
DR EMBL; AY165541; AAP21565.1; -.
DR HSP; P02568; ILCU.
DR FlyBase; FBgn0066214; Dnov\ActE2.
DR GO; GO:0005884; C:actin filament; IEA.
DR GO; GO:0003774; F:motor activity; IEA.
DR GO; GO:0005200; F:structural constituent of cytoskeleton; IEA.
DR InterPro; IPR004001; Actin.
DR Pfam; PF00022; Actin_1.
DR PRINTS; PR00190; ACTIN.
DR PROSITE; PS00406; ACTINS_1; 1.
KW Structural protein.
FT NON_TER 96
SQ SEQUENCE 96 AA; 10610 MW; 1454B7D9DD3F3564 CRC64;

Query Match 100.0%; Score 80; DB 2; Length 96;
Best Local Similarity 100.0%; Pred. No. 3.3e-06;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 FPSIVGRPRHQVMV 15
Db 32 FPSIVGRPRHQVMV 46

RESULT 11
Q9QZB4 PRELIMINARY; PRT; 99 AA.
ID Q9QZB4
AC Q9QZB4;
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE Cytoplasmic actin (Fragment).
OS Cavia porcellus (Guinea pig).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Hystricognathi; Caviidae; Cavia.
OX NCBI_TaxID=10141;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Adrenal gland;
RA Buelow H.E., Bernhard R.;
RL Submitted (OCT-1999) to the EMBL/GenBank/DBJ databases.
CC -!- FUNCTION: Actins are highly conserved proteins that are involved
in various types of cell motility and are ubiquitously expressed
in all eukaryotic cells (By similarity).
CC -!- SIMILARITY: Belongs to the actin family.
DR EMBL; AF191277; AAF13923.1; -.
DR HSP; P02577; INM1.
DR GO; GO:0005884; C:actin filament; IEA.
DR GO; GO:0003774; F:motor activity; IEA.
DR GO; GO:0005200; F:structural constituent of cytoskeleton; IEA.
DR InterPro; IPR004001; Actin.
DR InterPro; IPR004000; Actin_like.
DR Pfam; PF00022; Actin_1.
DR PRINTS; PR00190; ACTIN.
DR PROSITE; PS00406; ACTINS_1; 1.
KW Structural protein.
FT NON_TER 99
SQ SEQUENCE 99 AA; 10973 MW; 0562B645B8DAF17B CRC64;

Query Match 100.0%; Score 80; DB 2; Length 99;
Best Local Similarity 100.0%; Pred. No. 3.4e-06;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 FPSIVGRPRHQVMV 15
Db 32 FPSIVGRPRHQVMV 46

RESULT 10
Q86FV7 PRELIMINARY; PRT; 96 AA.
ID Q86FV7
AC Q86FV7;
DT 01-JUN-2003 (TrEMBLrel. 24, Created)
DT 01-JUN-2003 (TrEMBLrel. 24, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Actin E2 (Fragment)
GN Name=ActE2;
OS Drosophila americana (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
OX NCBI_TaxID=40366;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=15010-0951.0;
```



|||||

31 FPSIVGRPRHQGVW 45

# RESULT 12

Q86SB7 PRELIMINARY; PRT; 104 AA.  
 ID Q86SB7;  
 AC Q86SB7;  
 DT 01-JUN-2003 (TREMBLrel. 24, Created)  
 DT 01-JUN-2003 (TREMBLrel. 24, Last sequence update)  
 DT 01-MAR-2004 (TREMBLrel. 26, Last annotation update)  
 DE Actin (Fragment).  
 GN Name=ACT;  
 OS Gryllus bimaculatus (Two-spotted cricket).  
 CC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;  
 CC Neoptera; Orthopteroidea; Orthoptera; Ensifera; Gryllidae; Gryllinae;  
 CC Gryllus.  
 CC NCBI\_TaxID=6999;  
 RN [1]

SEQUENCE FROM N.A.

RP Zhang H., Shimyo Y., Hirose A., Mito T., Inoue Y., Ohuchi H.,  
 RA Loukeris T.G., Noji S.;  
 RT "Extrachromosomal transposition of the transposable element Minos  
 occurs in embryos of the cricket Gryllus bimaculatus."  
 RL Dev. Growth Differ. 0:0-0(2002).  
 CC -!- FUNCTION: Actins are highly conserved proteins that are involved  
 in various types of cell motility and are ubiquitously expressed  
 in all eukaryotic cells (By similarity).  
 CC -!- SIMILARITY: Belongs to the actin family.  
 CC EMBL; AB087882; BAC5093.1; -.  
 DR HSP; P02577; 1NM1.  
 DR GO; GO:0005884; C:actin filament; IEA.  
 DR GO; GO:0003774; F:motor activity; IEA.  
 DR GO; GO:0005200; F:structural constituent of cytoskeleton; IEA.  
 DR InterPro; IPR004001; Actin.  
 DR InterPro; IPR004000; Actin\_like.  
 DR Pfam; PF00022; Actin; 1.  
 DR PRINTS; PR00190; ACTIN.  
 DR PROSITE; PS00406; ACTINS\_1; 1.  
 KW Structural protein.  
 FT NON\_TER 104 104  
 SQ SEQUENCE 104 AA; 11511 MW; 332777A2B08669B0 CRC64;

Query Match 100.0%; Score 80; DB 2; Length 104;  
 Best Local Similarity 100.0%; Pred. No. 3.6e-06;  
 Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 FPSIVGRPRHQGVW 15  
 |||||

32 FPSIVGRPRHQGVW 46

# RESULT 13

O18548 PRELIMINARY; PRT; 108 AA.  
 ID O18548;  
 AC O18548;  
 DT 01-JAN-1998 (TREMBLrel. 05, Created)  
 DT 01-JAN-1998 (TREMBLrel. 05, Last sequence update)  
 DT 01-JUN-2003 (TREMBLrel. 24, Last annotation update)  
 DE Cytoplasmic actin Lvc4 (Fragment).  
 GN Name=Lvc4;  
 OS Lytechinus variegatus (Sea urchin).  
 CC Eukaryota; Metazoa; Echinodermata; Eleutherozoa; Echinozoa;  
 CC Echinoidea; Euechinoidea; Echinacea; Temnopleuroidea; Toxopneustidae;  
 CC Lytechinus.  
 CC NCBI\_TaxID=7654;  
 RN [1]

SEQUENCE FROM N.A.

RX MEDLINE=97333945; PubMed=9190067;  
 RA Kissinger J.C., Hahn J.H., Raff R.A.;  
 RT "Rapid evolution in a conserved gene family. Evolution of the actin  
 gene family in the sea urchin genus Helicodaris and related genera."  
 RL Mol. Biol. Evol. 14:654-665(1997).

CC -!- FUNCTION: Actins are highly conserved proteins that are involved  
 in various types of cell motility and are ubiquitously expressed  
 in all eukaryotic cells (By similarity).  
 CC -!- SIMILARITY: Belongs to the actin family.  
 CC EMBL; U82542; AAB66299.1; -.  
 DR HSP; P10983; 1D4X.  
 DR GO; GO:0005884; C:actin filament; IEA.  
 DR GO; GO:0003774; F:motor activity; IEA.  
 DR GO; GO:0005200; F:structural constituent of cytoskeleton; IEA.  
 DR InterPro; IPR004001; Actin.  
 DR InterPro; IPR004000; Actin\_like.  
 DR Pfam; PF00022; Actin; 1.  
 DR PRINTS; PR00190; ACTIN.  
 DR PROSITE; PS00406; ACTINS\_1; 1.  
 DR PROSITE; PS01132; ACTINS\_ACT\_LIKE; 1.  
 KW Structural protein.  
 FT NON\_TER 1 1  
 FT NON\_TER 108 108  
 SQ SEQUENCE 108 AA; 12094 MW; F670CB8A6CBD3AEA CRC64;

Query Match 100.0%; Score 80; DB 2; Length 108;  
 Best Local Similarity 100.0%; Pred. No. 3.8e-06;  
 Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 FPSIVGRPRHQGVW 15  
 |||||

18 FPSIVGRPRHQGVW 32

# RESULT 14

O18550 PRELIMINARY; PRT; 108 AA.  
 ID O18550;  
 AC O18550;  
 DT 01-JAN-1998 (TREMBLrel. 05, Created)  
 DT 01-JAN-1998 (TREMBLrel. 05, Last sequence update)  
 DT 01-JUN-2003 (TREMBLrel. 24, Last annotation update)  
 DE Cytoplasmic actin Lvc2 (Fragment).  
 GN Name=Lvc2;  
 OS Lytechinus variegatus (Sea urchin).  
 CC Eukaryota; Metazoa; Echinodermata; Eleutherozoa; Echinozoa;  
 CC Echinoidea; Euechinoidea; Echinacea; Temnopleuroidea; Toxopneustidae;  
 CC Lytechinus.  
 CC NCBI\_TaxID=7654;  
 RN [1]

SEQUENCE FROM N.A.

RP MEDLINE=97333945; PubMed=9190067;  
 RA Kissinger J.C., Hahn J.H., Raff R.A.;  
 RT "Rapid evolution in a conserved gene family. Evolution of the actin  
 gene family in the sea urchin genus Helicodaris and related genera."  
 RL Mol. Biol. Evol. 14:654-665(1997).

CC -!- FUNCTION: Actins are highly conserved proteins that are involved  
 in various types of cell motility and are ubiquitously expressed  
 in all eukaryotic cells (By similarity).  
 CC -!- SIMILARITY: Belongs to the actin family.  
 CC EMBL; U82544; AAB66301.1; -.  
 DR HSP; P10983; 1D4X.  
 DR GO; GO:0005884; C:actin filament; IEA.  
 DR GO; GO:0003774; F:motor activity; IEA.  
 DR GO; GO:0005200; F:structural constituent of cytoskeleton; IEA.  
 DR InterPro; IPR004001; Actin.  
 DR InterPro; IPR004000; Actin\_like.  
 DR Pfam; PF00022; Actin; 1.  
 DR PRINTS; PR00190; ACTIN.  
 DR PROSITE; PS00406; ACTINS\_1; 1.  
 DR PROSITE; PS01132; ACTINS\_ACT\_LIKE; 1.  
 KW Structural protein.  
 FT NON\_TER 1 1  
 FT NON\_TER 108 108  
 SQ SEQUENCE 108 AA; 12097 MW; F670D62101164AEA CRC64;

Query Match 100.0%; Score 80; DB 2; Length 108;  
 Best Local Similarity 100.0%; Pred. No. 3.8e-06;  
 Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 FPSIVGRPRHQGMV 15  
 |||||  
 Db 18 FPSIVGRPRHQGMV 32

RESULT 15  
 Q8JG71 PRELIMINARY; PRT; 121 AA.  
 AC Q8JG71;  
 DT 01-OCT-2002 (TrEMBLrel. 22, Created)  
 DT 01-OCT-2002 (TrEMBLrel. 22, Last sequence update)  
 DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)  
 DE Beta-actin (Fragment).  
 OS Oreochromis niloticus (Nile tilapia) (Tilapia nilotica).  
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;  
 OC Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;  
 OC Acanthomorpha; Acanthopterygii; Perciformes; Labroidae;  
 OC Cichlidae; Oreochromis.  
 OX NCBI\_TaxID=8128;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=22414672; PubMed=12527420; DOI=10.1016/S0167-4781(02)00534-1;  
 RA Hwang G.-L., Rahman M.A., Razak S.A., Schm F., Farahmand H., Smith A.,  
 RA Brooks C., Maclean N.;  
 RT "Isolation and characterization of tilapia beta-actin promoter and  
 RT comparison of its activity with carp beta-actin promoter.";  
 RL Biochim. Biophys. Acta 1625:11-18(2003).  
 CC -!- FUNCTION: Actins are highly conserved proteins that are involved  
 CC in various types of cell motility and are ubiquitously expressed  
 CC in all eukaryotic cells (By similarity).  
 CC -!- SIMILARITY: Belongs to the actin family.  
 DR EMBL; AY116536; AAM66751.1; -.  
 DR HSP; P10983; ID4X.  
 DR GO; GO:0005884; C:actin filament; IEA.  
 DR GO; GO:0003774; F:motor activity; IEA.  
 DR GO; GO:0005200; F:structural constituent of cytoskeleton; IEA.  
 DR InterPro; IPR004001; Actin.  
 DR InterPro; IPR004000; Actin\_like.  
 DR Pfam; PF00022; Actin; 1.  
 DR PRINTS; PR00190; ACTIN.  
 DR PROSITE; PS00406; ACTINS 1; 1.  
 DR PROSITE; PS01132; ACTINE\_ACT\_LIKE; 1.  
 KW Structural protein.  
 FT NON\_TER 121 121  
 SQ SEQUENCE 121 AA; 13458 MW; 04A58F271BBA0898 CRC64;

Query Match 100.0%; Score 80; DB 2; Length 121;  
 Best Local Similarity 100.0%; Pred. No. 4.2e-06;  
 Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 FPSIVGRPRHQGMV 15  
 |||||  
 Db 31 FPSIVGRPRHQGMV 45

Search completed: April 8, 2005, 12:03:06  
 Job time : 59.8667 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: April 8, 2005, 09:58:56 ; Search time 62.4 Seconds  
(without alignments)  
92.971 Million cell updates/sec

Title: US-09-423-351C-3

Perfect score: 82

Sequence: 1 GRPRHQGVGMGQK 15

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 2105692 seqs, 386760381 residues

Total number of hits satisfying chosen parameters: 2105692

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : A\_Geneseq\_16Dec04:\*

- 1: geneseqp1980s:\*
- 2: geneseqp1990s:\*
- 3: geneseqp2000s:\*
- 4: geneseqp2001s:\*
- 5: geneseqp2002s:\*
- 6: geneseqp2003as:\*
- 7: geneseqp2003bs:\*
- 8: geneseqp2004s:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	82	100.0	15	2 AAW92529	Beta-acti
2	82	100.0	19	2 AAW92543	Beta-acti
3	82	100.0	105	5 ABP42469	Human ova
4	82	100.0	157	4 ABB66853	Drosophil
5	82	100.0	201	5 ABP43176	Human ova
6	82	100.0	204	5 ABP43016	Human ova
7	82	100.0	310	8 ADR38362	piga3Cy3G
8	82	100.0	342	7 ADM05323	Human pro
9	82	100.0	374	3 AAB12385	Human bet
10	82	100.0	375	1 AAP61532	Sequence
11	82	100.0	375	2 AAR50328	Drug resi
12	82	100.0	375	3 AAB15017	Posttrans
13	82	100.0	375	3 AAY94569	Human car
14	82	100.0	375	3 AAB15016	Posttrans
15	82	100.0	375	5 ABB77395	Human act
16	82	100.0	375	6 ABR64271	Angiogene
17	82	100.0	375	6 ABM04830	Rat cytop
18	82	100.0	375	7 ADB85212	Rat actin
19	82	100.0	375	7 ADE61174	Rat prote
20	82	100.0	375	7 ADF30525	Rat angio
21	82	100.0	375	7 ADI63062	Human apo
22	82	100.0	375	7 ADI62970	Human apo
23	82	100.0	375	7 ADI63040	Human apo
24	82	100.0	375	8 ADI13002	Human ste
25	82	100.0	375	8 ADJ78489	Actin pro

26	82	100.0	375	8 ADL14103	Human ear
27	82	100.0	375	8 ADP04899	Sea squir
28	82	100.0	375	8 ADP12391	Protein e
29	82	100.0	375	8 ADQ26098	Gamma 1 a
30	82	100.0	375	8 ABO84772	Human can
31	82	100.0	375	8 ABM80841	Tumour-as
32	82	100.0	375	8 ADN23274	Bacterial
33	82	100.0	375	8 ADS88828	Amino aci
34	82	100.0	375	8 ADS88825	Amino aci
35	82	100.0	375	8 ADS88826	Amino aci
36	82	100.0	375	8 ADS88827	Amino aci
37	82	100.0	376	2 AAW19799	Gamma-smo
38	82	100.0	376	4 ABB58162	Drosophil
39	82	100.0	376	4 ABB61322	Drosophil
40	82	100.0	376	4 ABB60354	Drosophil
41	82	100.0	376	4 ABB64853	Drosophil
42	82	100.0	376	4 ABB63315	Drosophil
43	82	100.0	376	6 ABR62327	Pacific w
44	82	100.0	376	8 ADN03845	Antipsoi
45	82	100.0	377	3 AAB15014	Human ear

## ALIGNMENTS

RESULT 1  
AAW92529  
ID AAW92529 standard; peptide; 15 AA.  
XX AC AAW92529;  
XX AC  
DT 26-APR-1999 (first entry)  
XX DE Beta-actin reference peptide substrate #3.  
XX KW Peptide substrate; CCT; eukaryotic type II chaperonin complex; cyclin;  
KW binding agent; substrate-binding site; SBS; substrate folding; actin;  
KW tubulin; treatment; cancer; anticancer drug; viral infection; screening;  
KW reduced toxicity.  
XX OS Synthetic.  
XX PN WO9853322-A1.  
XX PD 26-NOV-1998.  
XX PF 22-MAY-1998; 98WO-GB001485.  
XX PR 23-MAY-1997; 97GB-00010762.  
XX PA (CANC-) INST CANCER RES ROYAL CANCER HOSPITAL.  
XX PI Willison K, Hynes G, Liou AK;  
XX DR WPI; 1999-070162/06.  
XX PT Identifying specific binding agents for substrate binding site in CCT  
PT chaperonin complex - also new peptide binding agents and their mimetics,  
PT and peptides containing a specific CCT binding site, used for treating  
cancer.  
XX PS Disclosure; Fig 10; 97pp; English.  
XX CC This invention describes a method which uses the CCT (eukaryotic type II  
CC chaperonin) complex or part of it, for identifying a binding agent that  
CC can occupy a substrate-binding site (SBS) on the CCT complex. By binding  
CC to the CCT complex, the binding agents block an SBS so that biological  
CC activity of the CCT complex is affected, particularly its ability to fold  
CC substrates such as actin, tubulin and cyclin. The binding agents are  
CC useful for treatment of cancer, particularly when used in combination  
CC with an anticancer drug, or viral infections. Nucleic acid fragments are  
CC used to screen for agents, e.g. binding agents that modulate interaction  
CC between the CCT complex and a protein that is to be folded. The binding

CC agents may target cells that are actively synthesising tubulin etc.  
CC (unlike known microtubule-stabilising agents that affect all cells), so  
CC should have reduced toxicity for normal cells. AAW92527-W92541 are  
CC peptide substrates used in the method of the invention  
XX  
SQ Sequence 15 AA;

Query Match 100.0%; Score 82; DB 2; Length 15;  
Best Local Similarity 100.0%; Pred. No. 6.8e-07;  
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GRPRHQGVVGMGQK 15  
DB 1 GRPRHQGVVGMGQK 15

RESULT 2  
AAW92543  
ID AAW92543 standard; peptide; 19 AA.

XX AAW92543;

XX 26-APR-1999 (first entry)

XX Beta-actin array biotinylated peptide substrate #1.

XX Peptide substrate; CCT; eukaryotic type II chaperonin complex; cyclin;  
KW binding agent; substrate-binding site; SBS; substrate folding; actin;  
KW tubulin; treatment; cancer; anticancer drug; viral infection; screening;  
KW reduced toxicity.

XX Synthetic.

XX Key Location/Qualifiers

FH Modified-site 1 /note= "Ser modified by biotin"

XX WO9853322-A1.

XX 26-NOV-1998.

XX 22-MAY-1998; 98WO-GB001485.

XX 23-MAY-1997; 97GB-00010762.

XX (CANC-) INST CANCER RES ROYAL CANCER HOSPITAL.

XX Willison K, Hynes G, Liou AK;

XX WPI; 1999-070162/06.

XX Identifying specific binding agents for substrate binding site in CCT  
PT chaperonin complex - also new peptide binding agents and their mimetics,  
PT and peptides containing a specific CCT binding site, used for treating  
PT cancer.

XX Disclosure; Fig 11; 97pp; English.

XX This invention describes a method which uses the CCT (eukaryotic type II  
CC chaperonin) complex or part of it, for identifying a binding agent that  
CC can occupy a substrate-binding site (SBS) on the CCT complex. By binding  
CC to the CCT complex, the binding agents block an SBS so that biological  
CC activity of the CCT complex is affected, particularly its ability to fold  
CC substrates such as actin, tubulin and cyclin. The binding agents are  
CC useful for treatment of cancer, particularly when used in combination  
CC with an anticancer drug, or viral infections. Nucleic acid fragments are  
CC used to screen for agents, e.g. binding agents that modulate interaction  
CC between the CCT complex and a protein that is to be folded. The binding  
CC agents may target cells that are actively synthesising tubulin etc.  
CC (unlike known microtubule-stabilising agents that affect all cells), so  
CC should have reduced toxicity for normal cells. AAW92543-W92549 are  
CC peptide substrates used in the method of the invention  
XX

SQ Sequence 19 AA;

Query Match 100.0%; Score 82; DB 2; Length 19;  
Best Local Similarity 100.0%; Pred. No. 8.7e-07;  
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GRPRHQGVVGMGQK 15  
DB 5 GRPRHQGVVGMGQK 19

RESULT 3  
ABP42469  
ID ABP42469 standard; protein; 105 AA.

XX ABP42469;

XX 22-AUG-2002 (first entry)

XX Human ovarian antigen HOCQ94, SEQ ID NO:3601.

XX Human; ovarian antigen; ovary; ovarian; breast; cancer; tumour;  
KW ovarian cancer; breast cancer; tumour; reproductive system disorder;  
KW infertility; pregnancy disorder; anovulation; polycystic ovary syndrome;  
KW PCOS; ovarian cyst; dysmenorrhea; endocrine disorder; infection;  
KW inflammatory condition; immune disorder; blood disorder;  
KW cardiovascular disorder; respiratory disorder; neurological disorder;  
KW gastrointestinal disorder; urinary system disorder; drug screening;  
KW gene therapy; chromosome mapping; forensic analysis;  
KW antibody preparation; cytostatic; immunomodulatory; neuroprotective;  
KW antiinflammatory; gynaecological; reproductive.

XX Homo sapiens.

XX WO200200677-A1.

XX 03-JAN-2002.

XX 07-JUN-2001; 2001WO-US018569.

XX 07-JUN-2000; 2000US-0209467P.

XX (HUMA-) HUMAN GENOME SCI INC.

XX Birse CE, Rosen CA;

XX WPI; 2002-147878/19.

XX N-PSDB; ABQ55546.

XX Isolated nucleic acid molecules encoding novel ovarian polypeptides,  
XX useful in the prevention, treatment and diagnosis of cancer (e.g. ovarian  
XX cancer), immune disorders, cardiovascular disorders and neurological  
XX diseases.

XX Claim 11; SEQ ID NO 3601; 2922pp; English.

XX The invention relates to 2175 novel human ovarian antigens (ABP41054-  
CC ABP43228) and to cDNAs encoding them (ABQ54131-ABQ56305), and also  
CC encompasses polypeptides 90% identical and polynucleotides 95% identical  
CC to the sequences of the invention. The invention additionally relates to  
CC recombinant vectors and host cells comprising human ovarian antigen  
CC polynucleotides, antibodies against human ovarian antigens, and the use  
CC of ovarian antigen polynucleotides and polypeptides in diagnosing,  
CC treating, prognosing or preventing various ovary and/or breast-related  
CC disorders. Such conditions include ovarian cancer and breast cancer, and  
CC metastatic tumours of ovarian or breast origin, reproductive system  
CC disorders (e.g., infertility, disorders of pregnancy, anovulation,  
CC polycystic ovary syndrome, ovarian cysts, and dysmenorrhoea), endocrine  
CC disorders, infections (e.g., chlamydia, HIV, toxoplasmosis, and toxic  
CC shock syndrome), inflammatory conditions (e.g., mastitis, oophoritis and  
CC vaginitis), immune disorders (e.g., congenital and acquired  
CC immunodeficiencies, autoimmune oophoritis, systemic lupus erythematosus),  
CC blood-related disorders (e.g., anaemia), cardiovascular disorders,  
XX

CC respiratory disorders, neurological disorders, gastrointestinal disorders  
 CC and urinary system disorders. Ovarian antigen polypeptides and  
 CC polynucleotides may also be used in screening for compounds which  
 CC modulate ovarian antigen expression or activity. The polynucleotides may  
 CC further be used for gene therapy, chromosome mapping, in the  
 CC identification of individuals and in forensic analysis, and the  
 CC polypeptides may be used as food additives or to prepare antibodies  
 CC useful in disease diagnosis, drug targeting and phenotyping. The present  
 CC sequence represents a human ovarian antigen of the invention. Note: The  
 CC sequence data for this patent did not form part of the printed  
 CC specification, but was obtained in electronic format directly from WIPO  
 CC at ftp.wipo.int/pub/published\_pct\_sequences

XX Sequence 105 AA;

Query Match 100.0%; Score 82; DB 5; Length 105;  
 Best Local Similarity 100.0%; Pred. No. 5.1e-06;  
 Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GRPRHQGVVMGQK 15  
 |||||  
 Db 37 GRPRHQGVVMGQK 51

RESULT 4

ABB66853  
 ID ABB66853 standard; protein; 157 AA.

AC ABB66853;

XX 26-MAR-2002 (first entry)

XX Drosophila melanogaster polypeptide SEQ ID NO 27351.

XX Drosophila; developmental biology; cell signalling; insecticide;  
 KW pharmaceutical.

XX Drosophila melanogaster.

XX WO200171042-A2.

XX 27-SEP-2001.

XX 23-MAR-2001; 2001WO-US009231.

XX 23-MAR-2000; 2000US-0191637P.

XX 11-JUL-2000; 2000US-00614150.

XX (PEKE ) PE CORP NY.

XX Venter JC, Adams M, Li PWD, Myers EW;

XX WPI; 2001-656860/75.

XX N-PSDB; ABL10956.

XX New isolated nucleic acid detection reagent for detecting 1000 or more  
 PT genes from Drosophila and for elucidating cell signaling and cell-cell  
 PT interactions.

XX Disclosure; SEQ ID NO 27351; 21pp + Sequence Listing; English.

XX The invention relates to an isolated nucleic acid detection reagent  
 CC capable of detecting 1000 or more genes from Drosophila. The invention is  
 CC useful in developmental biology and in elucidating cell signalling and  
 CC cell-cell interactions in higher eukaryotes for the development of  
 CC insecticides, therapeutics and pharmaceutical drugs. The invention  
 CC discloses genomic DNA sequences (ABLI16176-ABLI30511), expressed DNA  
 CC sequences (ABLI01840-ABLI16175) and the encoded proteins (ABBS57737-  
 CC ABB72072). The sequence data for this patent did not form part of the  
 CC printed specification, but was obtained in electronic format directly  
 CC from WIPO at ftp.wipo.int/pub/published\_pct\_sequences

XX Sequence 157 AA;

Query Match 100.0%; Score 82; DB 4; Length 157;  
 Best Local Similarity 100.0%; Pred. No. 7.7e-06;  
 Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GRPRHQGVVMGQK 15  
 |||||  
 Db 37 GRPRHQGVVMGQK 51

RESULT 5

ABP43176  
 ID ABP43176 standard; protein; 201 AA.

XX AC ABP43176;

XX 22-AUG-2002 (first entry)

XX Human ovarian antigen HVQCQ49, SEQ ID NO:4308.

XX Human; ovarian antigen; ovary; ovarian; breast; cancer; tumour;  
 KW ovarian cancer; breast cancer; tumour; reproductive system disorder;  
 KW infertility; pregnancy disorder; anovulation; polycystic ovary syndrome;  
 KW PCOS; ovarian cyst; dysmenorrhea; endocrine disorder; infection;  
 KW inflammatory condition; immune disorder; blood disorder;  
 KW cardiovascular disorder; respiratory disorder; neurological disorder;  
 KW gastrointestinal disorder; urinary system disorder; drug screening;  
 KW gene therapy; chromosome mapping; forensic analysis;  
 KW antibody preparation; cytostatic; immunomodulatory; neuroprotective;  
 KW antiinflammatory; gynaecological; reproductive.

XX Homo sapiens.

XX WO200200677-A1.

XX 03-JAN-2002.

XX 07-JUN-2001; 2001WO-US018569.

XX 07-JUN-2000; 2000US-0209467P.

XX (HUMA-) HUMAN GENOME SCI INC.

XX Birse CE, Rosen CA;

XX WPI; 2002-147878/19.

XX N-PSDB; ABQ56253.

XX Isolated nucleic acid molecules encoding novel ovarian polypeptides,  
 PT useful in the prevention, treatment and diagnosis of cancer (e.g. ovarian  
 PT cancer), immune disorders, cardiovascular disorders and neurological  
 PT diseases.

XX Claim 11; SEQ ID NO 4308; 2922pp; English.

XX The invention relates to 2175 novel human ovarian antigens (ABP41054-  
 CC ABP43228) and to cDNAs encoding them (ABQ54131-ABQ56305), and also  
 CC encompasses polypeptides 90% identical and polynucleotides 95% identical  
 CC to the sequences of the invention. The invention additionally relates to  
 CC recombinant vectors and host cells comprising human ovarian antigen  
 CC polynucleotides, antibodies against human ovarian antigens, and the use  
 CC of ovarian antigen polynucleotides and polypeptides in diagnosing,  
 CC treating, prognosing or preventing various ovary and/or breast-related  
 CC disorders. Such conditions include ovarian cancer and breast cancer, and  
 CC metastatic tumours of ovarian or breast origin, reproductive system  
 CC disorders (e.g., infertility, disorders of pregnancy, anovulation,  
 CC polycystic ovary syndrome, ovarian cysts, and dysmenorrhoea), endocrine  
 CC disorders, infections (e.g., chlamydia, HIV, toxoplasmosis, and toxic  
 CC shock syndrome), inflammatory conditions (e.g., mastitis, oophoritis and  
 CC vaginitis), immune disorders (e.g., congenital and acquired  
 CC immunodeficiencies, autoimmune oophoritis, systemic lupus erythematosus),  
 CC blood-related disorders (e.g., anaemia), cardiovascular disorders,  
 CC respiratory disorders, neurological disorders, gastrointestinal disorders

CC and urinary system disorders. Ovarian antigen polypeptides and  
CC polynucleotides may also be used in screening for compounds which  
CC modulate ovarian antigen expression or activity. The polynucleotides may  
CC further be used for gene therapy, chromosome mapping, in the  
CC identification of individuals and in forensic analysis, and the  
CC polypeptides may be used as food additives or to prepare antibodies  
CC useful in disease diagnosis, drug targeting and phenotyping. The present  
CC sequence represents a human ovarian antigen of the invention. Note: The  
CC sequence data for this patent did not form part of the printed  
CC specification, but was obtained in electronic format directly from WIPO  
CC at ftp.wipo.int/pub/published\_pct\_sequences  
XX Sequence 201 AA;  
SQ  
Query Match 100.0%; Score 82; DB 5; Length 201;  
Best Local Similarity 100.0%; Pred. No. 9.9e-06;  
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 GRPRHQGVWVGQK 15  
Db 53 GRPRHQGVWVGQK 67  
RESULT 6  
ID ABP43016 standard; protein; 204 AA.  
AC ABP43016;  
XX  
XX 22-AUG-2002 (first entry)  
DE Human ovarian antigen HSPB24, SEQ ID NO:4148.  
XX  
XX Human; ovarian antigen; ovary; ovarian; breast; cancer; tumour;  
KW ovarian cancer; breast cancer; tumour; reproductive system disorder;  
KW infertility; pregnancy disorder; anovulation; polycystic ovary syndrome;  
KW PCOS; ovarian cyst; dysmenorrhoea; endocrine disorder; infection;  
KW inflammatory condition; immune disorder; blood disorder;  
KW cardiovascular disorder; respiratory disorder; neurological disorder;  
KW gastrointestinal disorder; urinary system disorder; drug screening;  
KW gene therapy; chromosome mapping; forensic analysis;  
KW antibody preparation; cytostatic; immunomodulatory; neuroprotective;  
KW antiinflammatory; gynaecological; reproductive.  
XX  
XX Homo sapiens.  
OS  
XX WO200200677-A1.  
XX  
XX 03-JAN-2002.  
XX  
XX 07-JUN-2001; 2001WO-US018569.  
XX  
XX 07-JUN-2000; 2000US-0209467P.  
XX  
XX (HUMA-) HUMAN GENOME SCI INC.  
XX  
XX Birse CE, Rosen CA;  
PI  
XX WPI; 2002-147878/19.  
XX  
XX N-PSDB; ABQ56093.  
XX  
XX Isolated nucleic acid molecules encoding novel ovarian polypeptides,  
PT useful in the prevention, treatment and diagnosis of cancer (e.g. ovarian  
PT cancer), immune disorders, cardiovascular disorders and neurological  
PT diseases.  
XX  
XX Claim 11; SEQ ID NO 4148; 2922pp; English.  
XX  
XX The invention relates to 2175 novel human ovarian antigens (ABP41054-  
CC ABP43228) and to cDNAs encoding them (ABQ54131-ABQ56305), and also  
CC encompasses polypeptides 90% identical and polynucleotides 95% identical  
CC to the sequences of the invention. The invention additionally relates to  
CC recombinant vectors and host cells comprising human ovarian antigen

CC polynucleotides, antibodies against human ovarian antigens, and the use  
CC of ovarian antigen polynucleotides and polypeptides in diagnosing,  
CC treating, prognosing or preventing various ovary and/or breast-related  
CC disorders. Such conditions include ovarian cancer and breast cancer, and  
CC metastatic tumours of ovarian or breast origin, reproductive system  
CC disorders (e.g., infertility, disorders of pregnancy, anovulation,  
CC polycystic ovary syndrome, ovarian cysts, and dysmenorrhoea), endocrine  
CC disorders, infections (e.g., chlamydia, HIV, toxoplasmosis, and toxic  
CC shock syndrome), inflammatory conditions (e.g., mastitis, oophoritis and  
CC vaginitis), immune disorders (e.g., congenital and acquired  
CC immunodeficiencies, autoimmune oophoritis, systemic lupus erythematosus),  
CC blood-related disorders (e.g., anaemia), cardiovascular disorders,  
CC respiratory disorders, neurological disorders, gastrointestinal disorders  
CC and urinary system disorders. Ovarian antigen polypeptides and  
CC polynucleotides may also be used in screening for compounds which  
CC modulate ovarian antigen expression or activity. The polynucleotides may  
CC further be used for gene therapy, chromosome mapping, in the  
CC identification of individuals and in forensic analysis, and the  
CC polypeptides may be used as food additives or to prepare antibodies  
CC useful in disease diagnosis, drug targeting and phenotyping. The present  
CC sequence represents a human ovarian antigen of the invention. Note: The  
CC sequence data for this patent did not form part of the printed  
CC specification, but was obtained in electronic format directly from WIPO  
CC at ftp.wipo.int/pub/published\_pct\_sequences  
XX Sequence 204 AA;  
SQ  
Query Match 100.0%; Score 82; DB 5; Length 204;  
Best Local Similarity 100.0%; Pred. No. 1e-05;  
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 GRPRHQGVWVGQK 15  
Db 56 GRPRHQGVWVGQK 70  
RESULT 7  
ID ADR38362 standard; protein; 310 AA.  
XX ADR38362;  
AC ADR38362;  
XX  
XX 18-NOV-2004 (first entry)  
DT  
XX pigA3Cy3GFP plasmid protein product (Cycle3 GFP under an actin promoter).  
DE  
XX piggyBac; transposon; transposase; gene transfer; plasmid; Cycle3 GFP;  
KW actin promoter.  
KW  
XX Bombyx mori.  
OS  
XX Synthetic.  
OS  
XX Unidentified.  
OS  
XX Key Location/Qualifiers  
FH Misc-difference 20  
FT /note= "Encoded by GCC"  
XX  
XX JP2004236642-A.  
XX  
XX 26-AUG-2004.  
XX  
XX 10-FEB-2003; 2003JP-00032306.  
XX  
XX 10-FEB-2003; 2003JP-00032306.  
XX  
XX (TORA ) TORAY IND INC.  
XX  
XX WPI; 2004-608233/59.  
XX  
XX N-PSDB; ADR38361.  
XX  
XX Introducing foreign gene into insect cell, involves introducing DNA  
PT having foreign gene integrated between inverted repeat sequences of  
PT piggyBac transposon and piggyBac transposase that acts on inverted repeat

PT sequence.  
PS Disclosure; Page; 25pp; Japanese.  
XX  
CC This invention relates to a novel method for introducing a foreign gene  
CC into an insect cell. Specifically, it refers to a foreign gene occurring  
CC between inverted repeat sequences of a piggyBac transposon and using a  
CC piggyBac transposase for gene transfer to give a high recombinant  
CC efficiency. The present invention describes the insect cell as that of  
CC Lepidoptera insect origin, preferably it is the silk worm Bombyx mori,  
CC and stable integration occurs into the genome of this insect using the  
CC transposase enzyme. In particular, this gene transfer method uses the  
CC plasmid pIG3C3GFP which contains a Cyc1e3 GFP gene expressed under a  
CC silk worm actin promoter (A3 promoter) between a pair of inverted repeat  
CC sequences of the transposon piggyBac. This polypeptide sequence is the  
CC protein product from the pIG3C3GFP plasmid DNA that represents the  
CC silkworm actin promoter controlling expression of the Cyc1e3 GFP gene  
CC given in an exemplification of the invention. NOTE: This sequence is  
CC given as an embedded protein in the sequence listing and is not referred  
CC to further within the specification.  
XX  
SQ Sequence 310 AA;  
Query Match 100.0%; Score 82; DB 8; Length 310;  
Best Local Similarity 100.0%; Pred. No. 1.5e-05;  
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
Qy 1 GRPRHQGVWVGWGQK 15  
Db 37 GRPRHQGVWVGWGQK 51  
RESULT 8  
ADM05323  
ID ADM05323 standard; protein; 342 AA.  
XX  
AC ADM05323;  
XX  
DT 20-MAY-2004 (first entry)  
XX  
XX Human protein of the invention SEQ ID NO:4008.  
XX human; gene therapy; diagnostic marker; pharmaceutical.  
XX Homo sapiens.  
XX  
XX EPI347046-A1.  
XX  
XX PD 24-SEP-2003.  
XX  
XX PF 12-APR-2002; 2002EP-00008400.  
XX  
XX PR 22-MAR-2002; 2002JP-00137785.  
XX  
XX PA (REAS-) RES ASSOC BIOTECHNOLOGY.  
XX  
XX PI Isogai T, Sugiyama T, Otsuki T, Wakamatsu A, Sato H, Ishii S;  
XX PI Yamamoto J, Isono Y, Hio Y, Otsuka K, Nagai K, Tamechika I;  
XX PI Seki N, Yoshikawa T, Otsuka M, Nagahari K, Masuho Y;  
XX WPI; 2003-723558/69.  
XX  
XX DR N-PSDB; ADM02880.  
XX  
XX PT New polynucleotides and polypeptides are useful in gene therapy, for  
XX developing a diagnostic marker or medicines for regulating their  
XX expression and activity, or as a target of gene therapy.  
XX  
XX PS Claim 1; SEQ ID NO 4008; 305pp; English.  
XX  
XX CC The invention relates to a novel human polynucleotide and the encoded  
XX polypeptide. A polynucleotide of the invention may have a use in gene  
XX therapy. An oligonucleotide of the invention ADM06202-ADM06773 is useful  
XX as a primer for synthesizing the polynucleotide or as a probe for

CC detecting the polynucleotide. The polynucleotides ADM01316-ADM03758 are  
CC useful in gene therapy, for developing a diagnostic marker or medicines  
CC for regulating their expression and activity, or as a target of gene  
CC therapy. The proteins ADM03759-ADM06201 encoded by the polynucleotides  
CC are useful as pharmaceutical agents. The present sequence represents a  
CC protein sequence of the invention.  
XX  
SQ Sequence 342 AA;  
Query Match 100.0%; Score 82; DB 7; Length 342;  
Best Local Similarity 100.0%; Pred. No. 1.7e-05;  
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
Qy 1 GRPRHQGVWVGWGQK 15  
Db 38 GRPRHQGVWVGWGQK 52  
RESULT 9  
AAB12985  
ID AAB12985 standard; protein; 374 AA.  
XX  
XX AC AAB12985;  
XX  
XX DT 29-NOV-2000 (first entry)  
XX  
XX DE Human beta-actin protein sequence.  
XX  
XX KW Beta actin; post translational modification; sickle cell anaemia;  
XX KW irreversibly sickled cell; ISC; treatment.  
XX  
XX OS Homo sapiens.  
XX  
XX PN US6087398-A.  
XX  
XX PD 11-JUL-2000.  
XX  
XX PF 01-MAR-1996; 96US-00609236.  
XX  
XX PR 14-AUG-1995; 95US-0002288P.  
XX  
XX PA (SALA-) SOUTH ALABAMA MEDICAL SCI FOUND.  
XX  
XX PI Goodman SR;  
XX  
XX DR WPI; 2000-498210/44.  
XX  
XX PT Treatment for sickle cell anemia comprises administering a reducing agent  
XX PT to inhibit and reverse sickled cell formation in blood.  
XX  
XX PS Example 17; Fig 5B; 53pp; English.  
XX  
XX CC This invention provides a treatment for sickle cell anaemia, which  
XX comprises administering a reducing agent. The treatment inhibits the  
XX formation of irreversible sickled cells (ISC) and reverses ISC formation  
XX in the blood. The present sequence represents the human beta-actin  
XX protein. The sequence is used in the invention to demonstrate that a  
XX disulphide bridge is formed between cysteines 284 and 373 in ISC beta-  
XX actin as a post-translational modification. The reducing agent used in  
XX the treatment acts to correct this post-translational modification  
XX  
SQ Sequence 374 AA;  
Query Match 100.0%; Score 82; DB 3; Length 374;  
Best Local Similarity 100.0%; Pred. No. 1.9e-05;  
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
Qy 1 GRPRHQGVWVGWGQK 15  
Db 35 GRPRHQGVWVGWGQK 49  
RESULT 10

```

AAP61532
ID AAP61532 standard; protein; 375 AA.
XX
AC AAP61532;
XX
DT 25-MAR-2003 (revised)
DT 13-SEP-1991 (first entry)
XX
DE Sequence of beta-actin.
XX
KW Actin fibre; actin film; recombinant beta-actin.
XX
OS Homo sapiens.
XX
FN EPI74608-A.
XX
PD 19-MAR-1986.
XX
PF 05-SEP-1985; 85EP-00111225.
XX
PR 13-SEP-1984; 84US-00650958.
XX
PA (STRD ) UNIV LELAND STANFORD JUNIOR.
XX
PI Leavitt JC, Kedes LH, Gunning PW;
XX
DR WPI; 1986-077015/12.
DR N-PSDB; AAN60172.
XX
PT Beta-actin gene and regulatory elements - used for expression of
PT polypeptide(s) in mammalian host cells.
XX
PS Example; Page 23-24; 32pp; English.
XX
CC In the example, a beta-actin expression vector providing the beta- actin
CC promoter region, a polylinker and a polyadenylation signal was
CC constructed where the expression construct was present on a vector having
CC a bacterial origin of replication, as well as a marker for selection in a
CC mammalian host. (Updated on 25-MAR-2003 to correct PA field.)
XX
SQ Sequence 375 AA;
  Query Match          100.0%; Score 82; DB 1; Length 375;
  Best Local Similarity 100.0%; Pred. No. 1.9e-05;
  Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
  Qy 1 GRPRHQGVWVGQK 15
  Db 36 GRPRHQGVWVGQK 50

Beta-actin gene and regulatory elements - used for expression of
polypeptide(s) in mammalian host cells.
Example; Page 23-24; 32pp; English.
In the example, a beta-actin expression vector providing the beta- actin
promoter region, a polylinker and a polyadenylation signal was
constructed where the expression construct was present on a vector having
a bacterial origin of replication, as well as a marker for selection in a
mammalian host. (Updated on 25-MAR-2003 to correct PA field.)
Sequence 375 AA;
  Query Match          100.0%; Score 82; DB 1; Length 375;
  Best Local Similarity 100.0%; Pred. No. 1.9e-05;
  Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
  Qy 1 GRPRHQGVWVGQK 15
  Db 36 GRPRHQGVWVGQK 50

Result 11
AAR50328
ID AAR50328 standard; protein; 375 AA.
XX
AC AAR50328;
XX
DT 12-OCT-1994 (first entry)
XX
DE Drug resistant structural protein.
XX
KW Drug resistant; structural gene; expression vector; selective; marker;
KW transformation; thymidine-kinase deleted cell.
XX
OS Homo sapiens.
XX
FN JP06038773-A.
XX
PD 15-FEB-1994.
XX
PF 31-JAN-1992; 92JP-00045939.
XX
PR 31-JAN-1992; 92JP-00045939.

AAP61532
ID AAP61532 standard; protein; 375 AA.
XX
AC AAP61532;
XX
DT 25-MAR-2003 (revised)
DT 13-SEP-1991 (first entry)
XX
DE Sequence of beta-actin.
XX
KW Actin fibre; actin film; recombinant beta-actin.
XX
OS Homo sapiens.
XX
FN EPI74608-A.
XX
PD 19-MAR-1986.
XX
PF 05-SEP-1985; 85EP-00111225.
XX
PR 13-SEP-1984; 84US-00650958.
XX
PA (STRD ) UNIV LELAND STANFORD JUNIOR.
XX
PI Leavitt JC, Kedes LH, Gunning PW;
XX
DR WPI; 1986-077015/12.
DR N-PSDB; AAN60172.
XX
PT Beta-actin gene and regulatory elements - used for expression of
PT polypeptide(s) in mammalian host cells.
XX
PS Example; Page 23-24; 32pp; English.
XX
CC In the example, a beta-actin expression vector providing the beta- actin
CC promoter region, a polylinker and a polyadenylation signal was
CC constructed where the expression construct was present on a vector having
CC a bacterial origin of replication, as well as a marker for selection in a
CC mammalian host. (Updated on 25-MAR-2003 to correct PA field.)
XX
SQ Sequence 375 AA;
  Query Match          100.0%; Score 82; DB 2; Length 375;
  Best Local Similarity 100.0%; Pred. No. 1.9e-05;
  Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
  Qy 1 GRPRHQGVWVGQK 15
  Db 36 GRPRHQGVWVGQK 50

Result 12
AAB15017
ID AAB15017 standard; protein; 375 AA.
XX
AC AAB15017;
XX
DT 07-DEC-2000 (first entry)
XX
DE Posttranslationally modified human cardiac actin mutant E361G.
XX
KW Cardiac actin; ACTC; human; gene therapy; IDC; chromosome 15q14;
KW idiopathic dilated cardiomyopathy; mutant; mucin.
XX
OS Homo sapiens.
XX
OS Synthetic.
XX
FH Key Location/Qualifiers
FT Misc-difference 361 /note= "Wild-type Glu substituted by Gly"
FT
XX
XX US6063576-A.
XX
PD 16-MAY-2000.
XX
PF 29-JUN-1998; 98US-00106217.
XX
PR 29-JUN-1998; 98US-00106217.
XX
PA (UTAH ) UNIV UTAH RES POUND.
XX
PI Olson TM, Keating MT;
XX
DR WPI; 2000-375488/32.
DR N-PSDB; AAA73739.
XX
PT New cardiac actin gene comprising histidine to arginine or glycine to
PT glutamic acid substitution, useful in the diagnosis of diseases
PT associated with the mutation, specifically idiopathic dilated
PT cardiomyopathy.
XX
PS Claim 3; Page; 36pp; English.
XX
CC The present invention relates to human cardiac actin (ACTC). Genotypic

```



CC analyses show that ACTC is linked to idiopathic dilated cardiomyopathy (IDC). The ACTC gene maps to chromosome 15q14. Six PCR primer pairs (AAA49020 to AAA49031) have been developed to allow analysis of the six exons of ACTC. Studies of two families with IDC showed the presence of two mutations, one in exon 5 and one in exon 6. The present sequence is a mutant of posttranslationally modified ACTC without the initial two residues. The mutation is a Glu to Gly mutation at position 361. The ACTC protein may be used in a variety of methods for drug screening and for rational drug design. The ACTC gene may be used to treat IDC by gene therapy. Analysis of the ACTC gene provides early identification of subjects likely to develop or who already have IDC. Note: The present sequence is not shown in the specification but is derived from the ACTC sequence described in AAY94569

XX SQ Sequence 375 AA;

Query Match 100.0%; Score 82; DB 3; Length 375;  
Best Local Similarity 100.0%; Pred. No. 1.9e-05;  
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GRPRHQGVVGMGQK 15  
|||||  
Db 36 GRPRHQGVVGMGQK 50

RESULT 13  
AAY94569  
ID AAY94569 standard; protein; 375 AA.  
XX AC AAY94569;  
XX 07-DEC-2000 (first entry)  
XX Human cardiac actin protein after posttranslational modification.  
XX Cardiac actin; ACTC; human; gene therapy; IDC; 15q14;  
KW idiopathic dilated cardiomyopathy.  
XX Homo sapiens.  
XX US6063576-A.  
XX 16-MAY-2000.  
XX 29-JUN-1998; 98US-00106217.  
XX 29-JUN-1998; 98US-00106217.  
XX (UTAH ) UNIV UTAH RES FOUND.  
XX Olson TM, Keating MT;  
XX WPI; 2000-375488/32.  
XX N-PSDB; AAA49032.  
XX New cardiac actin gene comprising histidine to arginine or glycine to glutamic acid substitution, useful in the diagnosis of diseases associated with the mutation, specifically idiopathic dilated cardiomyopathy.  
XX Claim 3; Col 61-62; 36pp; English.  
XX The present sequence is human cardiac actin (ACTC) protein, without the initial Met-Cys which is posttranslationally removed. Genotypic analyses show that ACTC is linked to idiopathic dilated cardiomyopathy (IDC). The ACTC gene maps to chromosome 15q14. Six PCR primer pairs (AAA49020 to AAA49031) have been developed to allow analysis of the six exons of ACTC. Studies of two families with IDC showed the presence of two mutations, one in exon 5 and one in exon 6. The mutations are G to A in codon 312 (Arg312His) and A to G in codon 361 (Glu361Gly). The ACTC protein may be used in a variety of methods for drug screening and for rational drug design. The ACTC gene may be used to treat IDC by gene therapy. Analysis of the ACTC gene provides early identification of subjects likely to

CC develop or who already have IDC  
XX SQ Sequence 375 AA;

Query Match 100.0%; Score 82; DB 3; Length 375;  
Best Local Similarity 100.0%; Pred. No. 1.9e-05;  
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GRPRHQGVVGMGQK 15  
|||||  
Db 36 GRPRHQGVVGMGQK 50

RESULT 14  
AAB15016  
ID AAB15016 standard; protein; 375 AA.  
XX AC AAB15016;  
XX 07-DEC-2000 (first entry)  
XX Posttranslationally modified human cardiac actin mutant R312H.  
XX Cardiac actin; ACTC; human; gene therapy; IDC; chromosome 15q14;  
KW idiopathic dilated cardiomyopathy; mutant; mutein.  
XX Homo sapiens.  
XX Synthetic.  
XX Key Location/Qualifiers  
FH Misc-difference 312  
FT /note= "Wild-type Arg substituted by His"  
XX US6063576-A.  
XX 16-MAY-2000.  
XX 29-JUN-1998; 98US-00106217.  
XX 29-JUN-1998; 98US-00106217.  
XX (UTAH ) UNIV UTAH RES FOUND.  
XX Olson TM, Keating MT;  
XX WPI; 2000-375488/32.  
XX N-PSDB; AAA73738.  
XX New cardiac actin gene comprising histidine to arginine or glycine to glutamic acid substitution, useful in the diagnosis of diseases associated with the mutation, specifically idiopathic dilated cardiomyopathy.  
XX Claim 3; Page; 36pp; English.  
XX The present invention relates to human cardiac actin (ACTC). Genotypic analyses show that ACTC is linked to idiopathic dilated cardiomyopathy (IDC). The ACTC gene maps to chromosome 15q14. Six PCR primer pairs (AAA49020 to AAA49031) have been developed to allow analysis of the six exons of ACTC. Studies of two families with IDC showed the presence of two mutations, one in exon 5 and one in exon 6. The present sequence is a mutant of posttranslationally modified ACTC without the initial two residues. The mutation is a Arg to His mutation at position 312. The ACTC protein may be used in a variety of methods for drug screening and for rational drug design. The ACTC gene may be used to treat IDC by gene therapy. Analysis of the ACTC gene provides early identification of subjects likely to develop or who already have IDC. Note: The present sequence is not shown in the specification but is derived from the ACTC sequence described in AAY94569

XX SQ Sequence 375 AA;

Query Match 100.0%; Score 82; DB 3; Length 375;

Best Local Similarity 100.0%; Pred. No. 1.9e-05;  
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 1 GRPRHQGVVGVGQK 15  
Db 36 GRPRHQGVVGVGQK 50

RESULT 15

ABB77395  
ID ABB77395 standard; protein; 375 AA.

XX AC ABB77395;

DT 11-JUL-2002 (first entry)

DE Human actin gamma 1.

XX Human; dermatological; skin stress; ageing; spondin 2; cathepsin L;  
KW actin gamma 1; vimentin; fibroblast; skin; cosmetic; pharmaceutical.

XX OS Homo sapiens.

XX DE10050274-A1.

XX PD 18-APR-2002.

XX PF 09-OCT-2000; 2000DE-01050274.

XX PR 09-OCT-2000; 2000DE-01050274.

XX PA (HENK ) HENKEL KGAA.

XX PI Petersohn D, Schmitt G, Foerster T;

XX DR WPI; 2002-373046/41.

XX PT In vitro assays for skin stress and skin ageing includes determination of  
PT spondin 2, cathepsin L, actin gamma 1 and vimentin fragments secreted by  
XX skin fibroblasts.

XX PS Claim 6; Page 12; 14pp; German.

XX The invention relates to in vitro methods for the detection of skin  
CC stress and/or skin ageing in humans and animals based on the  
CC determination of spondin 2, cathepsin L, actin gamma 1 or vimentin  
CC fragments secreted by fibroblast from the skin under test. Use of the  
CC methods in a test for potential cosmetics and pharmaceuticals with an  
CC effect on these skin conditions and products containing vimentin  
CC fragments are also included. Products containing vimentin fragments are  
CC effective in the regulation, especially maintenance, of skin homeostasis

SQ Sequence 375 AA;

Query Match 100.0%; Score 82; DB 5; Length 375;  
Best Local Similarity 100.0%; Pred. No. 1.9e-05;  
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 1 GRPRHQGVVGVGQK 15  
Db 36 GRPRHQGVVGVGQK 50

Search completed: April 8, 2005, 10:50:42  
Job time : 62.4 secs

GenCore version 5.1.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: April 8, 2005, 09:58:57 ; Search time 17.5333 Seconds  
(without alignments)  
63.863 Million cell updates/sec

Title: US-09-423-351C-3  
Perfect score: 82  
Sequence: 1 GRPRHOGVMVGMGQK 15

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 513545 seqs, 74649064 residues

Total number of hits satisfying chosen parameters: 513545

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : Issued\_Patents\_AA:\*  
1: /cgn2\_6/prodata/1/1aa/5A-COMB.pep:\*  
2: /cgn2\_6/prodata/1/1aa/5B-COMB.pep:\*  
3: /cgn2\_6/prodata/1/1aa/6A-COMB.pep:\*  
4: /cgn2\_6/prodata/1/1aa/6B-COMB.pep:\*  
5: /cgn2\_6/prodata/1/1aa/PCTUS-COMB.pep:\*  
6: /cgn2\_6/prodata/1/1aa/backfiles1.pep:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	82	100.0	20	3	US-08-505-250-17
2	82	100.0	20	3	US-08-505-250-17
3	82	100.0	374	3	US-08-609-236-6
4	82	100.0	375	2	US-08-494-151-14
5	82	100.0	375	3	US-09-106-217-16
6	82	100.0	375	4	US-09-976-594-731
7	82	100.0	376	1	US-08-588-113-2
8	82	100.0	376	4	US-09-949-016-6100
9	82	100.0	377	3	US-09-106-217-2
10	82	100.0	377	4	US-09-919-172-33
11	82	100.0	377	4	US-09-917-254-53
12	82	100.0	386	4	US-09-949-016-7721
13	82	100.0	399	4	US-09-949-016-9424
14	82	100.0	402	4	US-09-949-016-10757
15	82	100.0	404	4	US-09-949-016-11313
16	81	98.8	377	4	US-09-248-796A-14109
17	77	93.9	71	4	US-09-621-976-6521
18	76	92.7	374	3	US-09-306-446C-2
19	75	91.5	375	3	US-09-171-337A-7
20	75	91.5	375	3	US-09-171-337A-8
21	75	91.5	375	4	US-09-631-022-7
22	75	91.5	375	4	US-09-631-022-8
23	47	57.3	397	4	US-09-902-540-9949
24	46	56.1	80	3	US-09-306-446C-4
25	44	53.7	318	3	US-09-199-637A-23
26	42	51.2	494	4	US-09-252-991A-32227
27	41.5	50.6	486	4	US-09-489-039A-10979

28	41	50.0	197	4	US-09-252-991A-32294	Sequence 32294, A
29	41	50.0	357	4	US-09-270-767-42609	Sequence 42609, A
30	41	50.0	376	4	US-09-538-092-1110	Sequence 1110, Ap
31	41	50.0	376	4	US-09-949-016-6656	Sequence 6656, Ap
32	41	50.0	376	4	US-09-949-016-8452	Sequence 8452, Ap
33	41	50.0	452	4	US-09-252-991A-23336	Sequence 23336, A
34	41	50.0	543	4	US-09-252-991A-18055	Sequence 18055, A
35	40	48.8	72	4	US-09-252-991A-28861	Sequence 28861, A
36	40	48.8	208	4	US-09-252-991A-19578	Sequence 19578, A
37	40	48.8	211	4	US-09-252-991A-18464	Sequence 18464, A
38	40	48.8	212	4	US-09-252-991A-29187	Sequence 29187, A
39	40	48.8	357	4	US-09-444-336-8	Sequence 8, Appli
40	40	48.8	357	4	US-09-444-336-9	Sequence 9, Appli
41	40	48.8	359	4	US-09-444-336-1	Sequence 1, Appli
42	40	48.8	362	4	US-09-444-336-2	Sequence 2, Appli
43	40	48.8	363	4	US-09-252-991A-27806	Sequence 27806, A
44	40	48.8	426	1	US-07-918-023-2	Sequence 2, Appli
45	40	48.8	599	4	US-09-949-016-8890	Sequence 8890, Ap

## ALIGNMENTS

RESULT 1  
US-08-505-250-17  
; Sequence 17, Application US/08505250  
; Patent No. 6183983  
; GENERAL INFORMATION:  
; APPLICANT: Sato, Haruya  
; APPLICANT: Yamamoto, Keiji  
; APPLICANT: Suzuki, Kokichi  
; APPLICANT: Ikeda, Masahiro  
; APPLICANT: Sakagami, Masahiro  
; APPLICANT: Taniguchi, Makoto  
; TITLE OF INVENTION: PROTEIN MODIFICATION METHOD  
; FILE REFERENCE: 110-511  
; CURRENT APPLICATION NUMBER: US/08/505,250  
; CURRENT FILING DATE: 1995-11-29  
; EARLIER APPLICATION NUMBER: PCT/JP95/00298  
; EARLIER FILING DATE: 1995-02-27  
; EARLIER APPLICATION NUMBER: JP 198187/94  
; EARLIER FILING DATE: 1994-08-23  
; NUMBER OF SEQ ID NOS: 53  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 17  
; LENGTH: 20  
; TYPE: PRT  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: Description of Artificial Sequence: synthetic  
; OTHER INFORMATION: peptide  
US-08-505-250-17

Query Match 100.0%; Score 82; DB 3; Length 20;  
Best Local Similarity 100.0%; Pred. No. 2.1e-07;  
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GRPRHOGVMVGMGQK 15  
Db 5 GRPRHOGVMVGMGQK 19

RESULT 2  
US-08-505-250-17  
; Sequence 17, Application US/08505250  
; Patent No. 6322996  
; GENERAL INFORMATION:  
; APPLICANT: Sato, Haruya  
; APPLICANT: Yamamoto, Keiji  
; APPLICANT: Suzuki, Kokichi  
; APPLICANT: Ikeda, Masahiro  
; APPLICANT: Sakagami, Masahiro  
; APPLICANT: Taniguchi, Makoto

## ; TITLE OF INVENTION: PROTEIN MODIFICATION METHOD

; FILE REFERENCE: 110-511  
; CURRENT APPLICATION NUMBER: US/08/505,250  
; CURRENT FILING DATE: 1995-11-29  
; PRIOR APPLICATION NUMBER: PCT/JP95/00298  
; PRIOR FILING DATE: 1995-02-27  
; PRIOR APPLICATION NUMBER: JP 198187/94  
; PRIOR FILING DATE: 1994-08-23  
; NUMBER OF SEQ ID NOS: 53  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 17  
; LENGTH: 20  
; TYPE: PRT

; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: Description of Artificial Sequence: synthetic  
; OTHER INFORMATION: peptide  
US-08-505-250-17

Query Match 100.0%; Score 82; DB 3; Length 20;  
Best Local Similarity 100.0%; Pred. No. 2.1e-07;  
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GRPRHQGVVMVGQK 15  
DB 5 GRPRHQGVVMVGQK 19

## RESULT 3

US-08-609-236-6  
; Sequence 6, Application US/08609236  
; Patent No. 6087398

; GENERAL INFORMATION:  
; APPLICANT: Steven R. Goodman  
; TITLE OF INVENTION: No. 6087398el Sickie Cell Anemia Treatment  
; NUMBER OF SEQUENCES: 6  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: McGregor & Adler, LLP  
; STREET: 8011 Candle Lane  
; CITY: Houston  
; STATE: Texas  
; COUNTRY: USA  
; ZIP: 77071

; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: Apple Macintosh  
; OPERATING SYSTEM: Macintosh  
; SOFTWARE: Microsoft Word for Macintosh  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/609,236  
; FILING DATE: March 1, 1996  
; CLASSIFICATION: 514

; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 60/002,288  
; FILING DATE: August 14, 1995  
; CLASSIFICATION: 514  
; ATTORNEY/AGENT INFORMATION:

; NAME: Benjamin Aaron Adler, Ph.D.  
; REGISTRATION NUMBER: 35,423  
; REFERENCE/DOCKET NUMBER: D5807  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 713-777-2321  
; TELEFAX: 713-777-6908

; INFORMATION FOR SEQ ID NO: 6:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 374  
; TYPE: Amino acid  
; STRANDEDNESS:

; TOPOLOGY: Linear  
; MOLECULE TYPE: Protein  
; DESCRIPTION: NO  
; HYPOTHETICAL: NO  
; ANTI-SENSE: NO

; FRAGMENT TYPE:  
; ORIGINAL SOURCE:

; STRAIN:  
; INDIVIDUAL ISOLATE:  
; DEVELOPMENTAL STAGE:

; TISSUE TYPE:  
; CELL TYPE:  
; CELL LINE:

US-08-609-236-6

Query Match 100.0%; Score 82; DB 3; Length 374;  
Best Local Similarity 100.0%; Pred. No. 4.8e-06;  
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GRPRHQGVVMVGQK 15  
DB 35 GRPRHQGVVMVGQK 49

## RESULT 4

US-08-494-151-14  
; Sequence 14, Application US/08494151  
; Patent No. 5840528

; GENERAL INFORMATION:  
; APPLICANT: Van Ooyen, Albert Johannes Joseph  
; TITLE OF INVENTION: Transformation of Phaffia rhodozyma  
; NUMBER OF SEQUENCES: 14  
; CORRESPONDENCE ADDRESS:

; ADDRESSEE: Morrison & Foerster  
; STREET: 2000 Pennsylvania Ave. N.W., Suite 5500  
; CITY: Washington, D.C.  
; COUNTRY: USA  
; ZIP: 20006-1812

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS

; SOFTWARE: PatentIn Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/494,151  
; FILING DATE: 23-JUN-1995

; CLASSIFICATION: 435  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Murashige, Kate H.  
; REGISTRATION NUMBER: 29,959  
; REFERENCE/DOCKET NUMBER: 24615-20039.01

; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (202) 887-1500  
; TELEFAX: (202) 887-0763

; TELEX: 90-4030

; INFORMATION FOR SEQ ID NO: 14:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 375 amino acids

; TYPE: amino acid

; TOPOLOGY: linear

; MOLECULE TYPE: protein

US-08-494-151-14

Query Match 100.0%; Score 82; DB 2; Length 375;  
Best Local Similarity 100.0%; Pred. No. 4.8e-06;  
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GRPRHQGVVMVGQK 15  
DB 36 GRPRHQGVVMVGQK 50

## RESULT 5

US-09-106-217-16  
; Sequence 16, Application US/09106217  
; Patent No. 6063576

; GENERAL INFORMATION:

; APPLICANT: Keating, Mark T.

APPLICANT: Olson, Timothy M.  
TITLE OF INVENTION: Actin Mutations in Dilated  
NUMBER OF SEQUENCES: 18  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Rothwell, Figg, Ernst & Kurz, P.C.  
STREET: 555 Thirteenth Street, N.W., Suite 701 East  
CITY: Washington  
STATE: DC  
COUNTRY: U.S.A.  
ZIP: 20004

COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/106,217  
FILING DATE:

CLASSIFICATION:  
ATTORNEY/AGENT INFORMATION:  
NAME: Saxe, Stephen A.  
REGISTRATION NUMBER: 38,609  
REFERENCE/DOCKET NUMBER: 2323-125  
TELEPHONE: 202-783-6040  
TELEFAX: 202-783-6031  
INFORMATION FOR SEQ ID NO: 16:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 375 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein

US-09-106-217-16  
Query Match 100.0%; Score 82; DB 3; Length 375;  
Best Local Similarity 100.0%; Pred. No. 4.8e-06;  
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Yy 1 GRPRHQGVVMGQK 15  
Db 36 GRPRHQGVVMGQK 50

RESULT 6  
US-09-976-594-731  
Sequence 731, Application US/09976594  
Patent No. 6673549  
GENERAL INFORMATION:  
APPLICANT: Furness, Michael  
TITLE OF INVENTION: GENES EXPRESSED IN C3A LIVER CELL CULTURES TREATED WITH STEROIDS  
FILE REFERENCE: PA-0041 US  
CURRENT APPLICATION NUMBER: US/09/976,594  
CURRENT FILING DATE: 2001-10-12  
PRIOR APPLICATION NUMBER: 60/240,409  
PRIOR FILING DATE: 2000-10-12  
NUMBER OF SEQ ID NOS: 1143  
SOFTWARE: PERL Program  
SEQ ID NO 731  
LENGTH: 375  
TYPE: PRT  
ORGANISM: Homo sapiens  
FEATURE:  
NAME/KEY: misc feature  
OTHER INFORMATION: Incyte ID No. 6673549 1837317CD1

US-09-976-594-731  
Query Match 100.0%; Score 82; DB 4; Length 375;  
Best Local Similarity 100.0%; Pred. No. 4.8e-06;  
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Yy 1 GRPRHQGVVMGQK 15  
Db 36 GRPRHQGVVMGQK 50

RESULT 7  
US-09-588-113-2  
Sequence 113-2, Application US/08588113  
Patent No. 5710003  
GENERAL INFORMATION:  
APPLICANT: McHugh, Kirk M.  
TITLE OF INVENTION: DIAGNOSTIC TEST FOR DETERMINING  
MALIGNANCY OF SMOOTH MUSCLE TUMORS  
NUMBER OF SEQUENCES: 4  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Woodcock Washburn Kurtz Mackiewicz &  
STREET: One Liberty Place - 46th Floor  
CITY: Philadelphia  
STATE: PA  
COUNTRY: USA  
ZIP: 19103  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/588,113  
FILING DATE:  
CLASSIFICATION: 436  
ATTORNEY/AGENT INFORMATION:  
NAME: Ralph, Rebecca L.  
REGISTRATION NUMBER: 35,152  
REFERENCE/DOCKET NUMBER: T0U-1652  
TELEPHONE: 215-568-3100  
TELEFAX: 215-568-3439  
INFORMATION FOR SEQ ID NO: 2:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 376 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein

US-08-588-113-2  
Query Match 100.0%; Score 82; DB 1; Length 376;  
Best Local Similarity 100.0%; Pred. No. 4.8e-06;  
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Yy 1 GRPRHQGVVMGQK 15  
Db 37 GRPRHQGVVMGQK 51

RESULT 8  
US-09-949-016-6100  
Sequence 6100, Application US/09949016  
Patent No. 6812339  
GENERAL INFORMATION:  
APPLICANT: VENTER, J. Craig et al.  
TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED  
WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF  
FILE REFERENCE: CL001307  
CURRENT APPLICATION NUMBER: US/09/949,016  
CURRENT FILING DATE: 2000-04-14  
PRIOR APPLICATION NUMBER: 60/241,755  
PRIOR FILING DATE: 2000-10-20  
PRIOR APPLICATION NUMBER: 60/237,768  
PRIOR FILING DATE: 2000-10-03  
PRIOR APPLICATION NUMBER: 60/231,498  
PRIOR FILING DATE: 2000-09-08  
NUMBER OF SEQ ID NOS: 207012

US-09-949-016-6100  
Query Match 100.0%; Score 82; DB 4; Length 375;  
Best Local Similarity 100.0%; Pred. No. 4.8e-06;  
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Yy 1 GRPRHQGVVMGQK 15  
Db 37 GRPRHQGVVMGQK 51

```
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 6100
; LENGTH: 376
; TYPE: PRT
; ORGANISM: Human
US-09-949-016-6100

Query Match      100.0%; Score 82; DB 4; Length 376;
Best Local Similarity 100.0%; Pred. No. 4.8e-06;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      1 GRPRHQGVVMVGQK 15
Db      37 GRPRHQGVVMVGQK 51

RESULT 9
US-09-106-217-2
; Sequence 2, Application US/09106217
; Patent No. 6063576
; GENERAL INFORMATION:
; APPLICANT: Keating, Mark T.
; APPLICANT: Olson, Timothy M.
; TITLE OF INVENTION: Actin Mutations in Dilated
; TITLE OF INVENTION: Cardiomyopathy, a Heritable Form of Heart Failure
; NUMBER OF SEQUENCES: 18
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Rothwell, Figg, Ernst & Kurz, P.C.
; STREET: 555 Thirteenth Street, N.W., Suite 701 East
; CITY: Washington
; STATE: DC
; COUNTRY: U.S.A.
; ZIP: 20004
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/106,217
; FILING DATE:
; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
; NAME: Saxe, Stephen A.
; REGISTRATION NUMBER: 38,609
; REFERENCE/DOCKET NUMBER: 2323-125
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 202-783-6040
; TELEFAX: 202-783-6031
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 377 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-09-106-217-2

Query Match      100.0%; Score 82; DB 3; Length 377;
Best Local Similarity 100.0%; Pred. No. 4.8e-06;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      1 GRPRHQGVVMVGQK 15
Db      38 GRPRHQGVVMVGQK 52

RESULT 10
US-09-919-172-33
; Sequence 33, Application US/09919172
; Patent No. 6673545
; GENERAL INFORMATION:
; APPLICANT: Paris, Mary
```

```
; APPLICANT: Turner, Christopher M.
; TITLE OF INVENTION: PROSTATE CANCER MARKERS
; FILE REFERENCE: PA-0036 US
; CURRENT APPLICATION NUMBER: US/09/919,172
; CURRENT FILING DATE: 2001-07-30
; PRIOR APPLICATION NUMBER: 60/222,469
; PRIOR FILING DATE: 2000-07-28
; NUMBER OF SEQ ID NOS: 102
; SOFTWARE: PERL Program
; SEQ ID NO 33
; LENGTH: 377
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; OTHER INFORMATION: Incyte ID No. 6673545 1709118CD1
US-09-919-172-33

Query Match      100.0%; Score 82; DB 4; Length 377;
Best Local Similarity 100.0%; Pred. No. 4.8e-06;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      1 GRPRHQGVVMVGQK 15
Db      38 GRPRHQGVVMVGQK 52

RESULT 11
US-09-917-254-53
; Sequence 53, Application US/09917254
; Patent No. 6703204
; GENERAL INFORMATION:
; APPLICANT: Mutter, George
; APPLICANT: Baak, Jan
; TITLE OF INVENTION: Prognostic Classification of Breast Cancer
; FILE REFERENCE: B0801/7224(JRV)
; CURRENT APPLICATION NUMBER: US/09/917,254
; CURRENT FILING DATE: 2001-07-27
; PRIOR APPLICATION NUMBER: US 60/222,093
; PRIOR FILING DATE: 2000-07-28
; NUMBER OF SEQ ID NOS: 102
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 53
; LENGTH: 377
; TYPE: PRT
; ORGANISM: Homo Sapiens
US-09-917-254-53

Query Match      100.0%; Score 82; DB 4; Length 377;
Best Local Similarity 100.0%; Pred. No. 4.8e-06;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      1 GRPRHQGVVMVGQK 15
Db      38 GRPRHQGVVMVGQK 52

RESULT 12
US-09-949-016-7721
; Sequence 7721, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; TITLE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CLO01307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
```

; PRIOR FILING DATE: 2000-09-08  
; NUMBER OF SEQ ID NOS: 207012  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 7721  
; LENGTH: 386  
; TYPE: PRT  
; ORGANISM: Human  
US-09-949-016-7721

Query Match 100.0%; Score 82; DB 4; Length 386;  
Best Local Similarity 100.0%; Pred. No. 5e-06;  
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GRPRHQGVVVGWGQK 15  
Db 47 GRPRHQGVVVGWGQK 61

RESULT 13  
US-09-949-016-9424  
; Sequence 9424, Application US/09949016  
; Patent No. 6812339  
; GENERAL INFORMATION:  
; APPLICANT: VENTER, J. Craig et al.  
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED  
; FILE REFERENCE: CL001307  
; CURRENT APPLICATION NUMBER: US/09/949,016  
; CURRENT FILING DATE: 2000-04-14  
; PRIOR APPLICATION NUMBER: 60/241,755  
; PRIOR FILING DATE: 2000-10-20  
; PRIOR APPLICATION NUMBER: 60/237,768  
; PRIOR FILING DATE: 2000-10-03  
; PRIOR APPLICATION NUMBER: 60/231,498  
; PRIOR FILING DATE: 2000-09-08  
; NUMBER OF SEQ ID NOS: 207012  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 9424  
; LENGTH: 399  
; TYPE: PRT  
; ORGANISM: Human  
US-09-949-016-9424

Query Match 100.0%; Score 82; DB 4; Length 399;  
Best Local Similarity 100.0%; Pred. No. 5.1e-06;  
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GRPRHQGVVVGWGQK 15  
Db 60 GRPRHQGVVVGWGQK 74

RESULT 14  
US-09-949-016-10757  
; Sequence 10757, Application US/09949016  
; Patent No. 6812339  
; GENERAL INFORMATION:  
; APPLICANT: VENTER, J. Craig et al.  
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED  
; FILE REFERENCE: CL001307  
; CURRENT APPLICATION NUMBER: US/09/949,016  
; CURRENT FILING DATE: 2000-04-14  
; PRIOR APPLICATION NUMBER: 60/241,755  
; PRIOR FILING DATE: 2000-10-20  
; PRIOR APPLICATION NUMBER: 60/237,768  
; PRIOR FILING DATE: 2000-10-03  
; PRIOR APPLICATION NUMBER: 60/231,498  
; PRIOR FILING DATE: 2000-09-08  
; NUMBER OF SEQ ID NOS: 207012  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 10757  
; LENGTH: 402

; TYPE: PRT  
; ORGANISM: Human  
US-09-949-016-10757

Query Match 100.0%; Score 82; DB 4; Length 402;  
Best Local Similarity 100.0%; Pred. No. 5.2e-06;  
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GRPRHQGVVVGWGQK 15  
Db 63 GRPRHQGVVVGWGQK 77

RESULT 15  
US-09-949-016-11313  
; Sequence 11313, Application US/09949016  
; Patent No. 6812339  
; GENERAL INFORMATION:  
; APPLICANT: VENTER, J. Craig et al.  
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED  
; FILE REFERENCE: CL001307  
; CURRENT APPLICATION NUMBER: US/09/949,016  
; CURRENT FILING DATE: 2000-04-14  
; PRIOR APPLICATION NUMBER: 60/241,755  
; PRIOR FILING DATE: 2000-10-20  
; PRIOR APPLICATION NUMBER: 60/237,768  
; PRIOR FILING DATE: 2000-10-03  
; PRIOR APPLICATION NUMBER: 60/231,498  
; PRIOR FILING DATE: 2000-09-08  
; NUMBER OF SEQ ID NOS: 207012  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 11313  
; LENGTH: 404  
; TYPE: PRT  
; ORGANISM: Human  
US-09-949-016-11313

Query Match 100.0%; Score 82; DB 4; Length 404;  
Best Local Similarity 100.0%; Pred. No. 5.2e-06;  
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GRPRHQGVVVGWGQK 15  
Db 65 GRPRHQGVVVGWGQK 79

Search completed: April 8, 2005, 12:07:34  
Job time : 17.5333 secs

**THIS PAGE BLANK (USPTO)**



GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: April 8, 2005, 10:53:18 ; Search time 42.6667 Seconds  
(without alignments)  
116.718 Million cell updates/sec

Title: US-09-423-351C-3

Perfect score: 82

Sequence: 1 GRPRHQGVWVGWQK 15

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1418010 seqs, 331997259 residues

Total number of hits satisfying chosen parameters: 1418010

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications AA:\*

- 1: /cgn2\_6/ptodata/1/pubpaa/US07\_PUBCOMB.pep.\*
- 2: /cgn2\_6/ptodata/1/pubpaa/PCT\_NEW\_PUB.pep.\*
- 3: /cgn2\_6/ptodata/1/pubpaa/US05\_NEW\_PUB.pep.\*
- 4: /cgn2\_6/ptodata/1/pubpaa/US06\_PUBCOMB.pep.\*
- 5: /cgn2\_6/ptodata/1/pubpaa/US07\_NEW\_PUB.pep.\*
- 6: /cgn2\_6/ptodata/1/pubpaa/PCTUS\_PUBCOMB.pep.\*
- 7: /cgn2\_6/ptodata/1/pubpaa/US08\_NEW\_PUB.pep.\*
- 8: /cgn2\_6/ptodata/1/pubpaa/US08\_PUBCOMB.pep.\*
- 9: /cgn2\_6/ptodata/1/pubpaa/US09A\_PUBCOMB.pep.\*
- 10: /cgn2\_6/ptodata/1/pubpaa/US09B\_PUBCOMB.pep.\*
- 11: /cgn2\_6/ptodata/1/pubpaa/US09C\_PUBCOMB.pep.\*
- 12: /cgn2\_6/ptodata/1/pubpaa/US09\_NEW\_PUB.pep.\*
- 13: /cgn2\_6/ptodata/1/pubpaa/US10A\_PUBCOMB.pep.\*
- 14: /cgn2\_6/ptodata/1/pubpaa/US10B\_PUBCOMB.pep.\*
- 15: /cgn2\_6/ptodata/1/pubpaa/US10C\_PUBCOMB.pep.\*
- 16: /cgn2\_6/ptodata/1/pubpaa/US10D\_PUBCOMB.pep.\*
- 17: /cgn2\_6/ptodata/1/pubpaa/US10\_NEW\_PUB.pep.\*
- 18: /cgn2\_6/ptodata/1/pubpaa/US11\_NEW\_PUB.pep.\*
- 19: /cgn2\_6/ptodata/1/pubpaa/US60\_NEW\_PUB.pep.\*
- 20: /cgn2\_6/ptodata/1/pubpaa/US60\_PUBCOMB.pep.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Match	Length	ID	Description
1	82	100.0	105	15 US-10-264-049-3601	Sequence 3601, Ap
2	82	100.0	142	15 US-10-424-599-184471	Sequence 184471, Ap
3	82	100.0	201	15 US-10-264-049-4308	Sequence 4308, Ap
4	82	100.0	204	15 US-10-264-049-4148	Sequence 4148, Ap
5	82	100.0	219	14 US-10-002-631C-56	Sequence 56, Appl
6	82	100.0	342	15 US-10-108-260A-4008	Sequence 4008, Ap
7	82	100.0	371	16 US-10-322-281-68	Sequence 68, Appl
8	82	100.0	375	14 US-10-205-194-93	Sequence 93, Appl
9	82	100.0	375	14 US-10-316-253-88	Sequence 88, Appl
10	82	100.0	375	15 US-10-369-493-5927	Sequence 5927, Ap
11	82	100.0	375	15 US-10-205-331-94	Sequence 94, Appl
12	82	100.0	375	15 US-10-260-708-82	Sequence 82, Appl
13	82	100.0	376	15 US-10-341-434-101	Sequence 101, Appl

14	82	100.0	377	9 US-09-919-172-33	Sequence 33, Appl
15	82	100.0	377	15 US-10-236-031B-64	Sequence 64, Appl
16	82	100.0	377	15 US-10-333-680-4	Sequence 4, Appl
17	82	100.0	398	16 US-10-322-281-63	Sequence 63, Appl
18	82	100.0	413	9 US-09-925-301-1436	Sequence 1436, Ap
19	81	98.8	375	15 US-10-369-493-1786	Sequence 1786, Ap
20	81	98.8	375	16 US-10-477-369-1	Sequence 1, Appl
21	77	93.9	219	14 US-10-029-386-32056	Sequence 32056, A
22	77	93.9	393	15 US-10-092-900A-310	Sequence 310, App
23	76	92.7	60	15 US-10-424-599-161246	Sequence 161246, A
24	76	92.7	108	16 US-10-767-701-35226	Sequence 35226, A
25	76	92.7	135	16 US-10-767-701-47318	Sequence 47318, A
26	76	92.7	136	15 US-10-424-599-280811	Sequence 280811, A
27	76	92.7	197	15 US-10-424-599-280810	Sequence 280810, A
28	76	92.7	208	15 US-10-424-599-282905	Sequence 282905, A
29	76	92.7	230	16 US-10-767-701-42770	Sequence 42770, A
30	76	92.7	248	15 US-10-424-599-162684	Sequence 162684, A
31	76	92.7	347	15 US-10-424-599-283333	Sequence 283333, A
32	76	92.7	375	15 US-10-369-493-2436	Sequence 2436, Ap
33	76	92.7	376	16 US-10-437-963-179917	Sequence 179917, A
34	76	92.7	376	16 US-10-437-963-202420	Sequence 202420, A
35	76	92.7	376	16 US-10-767-701-45677	Sequence 45677, A
36	76	92.7	376	16 US-10-767-701-46090	Sequence 46090, A
37	76	92.7	377	14 US-10-338-777-52	Sequence 52, Appl
38	76	92.7	377	15 US-10-424-599-162009	Sequence 162009, A
39	76	92.7	377	15 US-10-424-599-162685	Sequence 162685, A
40	76	92.7	377	15 US-10-424-599-162686	Sequence 162686, A
41	76	92.7	377	15 US-10-424-599-222816	Sequence 222816, A
42	76	92.7	377	15 US-10-424-599-223492	Sequence 223492, A
43	76	92.7	377	15 US-10-424-599-283336	Sequence 283336, A
44	76	92.7	377	15 US-10-425-114-66124	Sequence 66124, A
45	76	92.7	377	16 US-10-437-963-121952	Sequence 121952, A

#### ALIGNMENTS

RESULT 1  
US-10-264-049-3601  
; Sequence 3601, Application US/10264049  
; Publication No. US20040005579A1  
; GENERAL INFORMATION:  
; APPLICANT: Birse et al.  
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies  
; FILE REFERENCE: PA13321  
; CURRENT APPLICATION NUMBER: US/10/264,049  
; CURRENT FILING DATE: 2002-10-04  
; PRIOR APPLICATION NUMBER: PCT/US01/18569  
; PRIOR FILING DATE: 2001-06-07  
; PRIOR APPLICATION NUMBER: US 60/209,467  
; PRIOR FILING DATE: 2000-06-07  
; NUMBER OF SEQ ID NOS: 4360  
; SOFTWARE: PatentIn Ver. 3.1  
; SEQ ID NO 3601  
; LENGTH: 105  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
; FEATURE:  
; NAME/KEY: MISC FEATURE  
; LOCATION: (103)  
; OTHER INFORMATION: Xaa equals any of the twenty naturally occurring L-amino acids  
US-10-264-049-3601

Query Match 100.0%; Score 82; DB 15; Length 105;  
Best Local Similarity 100.0%; Pred. No. 5e-06;  
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GRPRHQGVWVGWQK 15  
DB 37 GRPRHQGVWVGWQK 51

RESULT 2

US-10-424-599-184471  
; Sequence 184471, Application US/10424599  
; Publication No. US20040031072A1  
; GENERAL INFORMATION:  
; APPLICANT: La Rosa Thomas J  
; APPLICANT: Kovalic David K  
; APPLICANT: Zhou Yihua  
; APPLICANT: Cao Yongwei  
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With  
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement  
; FILE REFERENCE: 38-21(53223)B  
; CURRENT APPLICATION NUMBER: US/10/424,599  
; CURRENT FILING DATE: 2003-04-28  
; NUMBER OF SEQ ID NOS: 285684  
; SEQ ID NO 184471  
; LENGTH: 142  
; TYPE: PRT  
; ORGANISM: Glycine max  
; FEATURE:  
; OTHER INFORMATION: Clone ID: PAT\_MRT3847\_137594C.1.pap  
US-10-424-599-184471

Query Match 100.0%; Score 82; DB 15; Length 142;  
Best Local Similarity 100.0%; Pred. No. 6.8e-06;  
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GRPRHQGVVMGQK 15  
Db 45 GRPRHQGVVMGQK 59  
|||||

## RESULT 3

US-10-264-049-4308  
; Sequence 4308, Application US/10264049  
; Publication No. US20040005579A1  
; GENERAL INFORMATION:  
; APPLICANT: Birse et al.  
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies  
; FILE REFERENCE: PA133P1  
; CURRENT APPLICATION NUMBER: US/10/264,049  
; CURRENT FILING DATE: 2002-10-04  
; PRIOR APPLICATION NUMBER: PCT/US01/18569  
; PRIOR FILING DATE: 2001-06-07  
; PRIOR APPLICATION NUMBER: US 60/209,467  
; PRIOR FILING DATE: 2000-06-07  
; NUMBER OF SEQ ID NOS: 4360  
; SOFTWARE: PatentIn Ver. 3.1  
; SEQ ID NO 4308  
; LENGTH: 201  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
; FEATURE:  
; NAME/KEY: MISC FEATURE  
; LOCATION: (135)  
; OTHER INFORMATION: Xaa equals any of the twenty naturally occurring L-amino acids  
; FEATURE:  
; NAME/KEY: MISC FEATURE  
; LOCATION: (136)  
; OTHER INFORMATION: Xaa equals any of the twenty naturally occurring L-amino acids  
; FEATURE:  
; NAME/KEY: MISC FEATURE  
; LOCATION: (140)  
; OTHER INFORMATION: Xaa equals any of the twenty naturally occurring L-amino acids  
; FEATURE:  
; NAME/KEY: MISC FEATURE  
; LOCATION: (143)  
; OTHER INFORMATION: Xaa equals any of the twenty naturally occurring L-amino acids  
; FEATURE:  
; NAME/KEY: MISC FEATURE  
; LOCATION: (144)  
; OTHER INFORMATION: Xaa equals any of the twenty naturally occurring L-amino acids  
; FEATURE:  
; NAME/KEY: MISC FEATURE

; LOCATION: (145)  
; OTHER INFORMATION: Xaa equals any of the twenty naturally occurring L-amino acids  
; FEATURE:  
; NAME/KEY: MISC FEATURE  
; LOCATION: (151)  
; OTHER INFORMATION: Xaa equals any of the twenty naturally occurring L-amino acids  
; FEATURE:  
; NAME/KEY: MISC FEATURE  
; LOCATION: (159)  
; OTHER INFORMATION: Xaa equals any of the twenty naturally occurring L-amino acids  
; FEATURE:  
; NAME/KEY: MISC FEATURE  
; LOCATION: (172)  
; OTHER INFORMATION: Xaa equals any of the twenty naturally occurring L-amino acids  
; FEATURE:  
; NAME/KEY: MISC FEATURE  
; LOCATION: (174)  
; OTHER INFORMATION: Xaa equals any of the twenty naturally occurring L-amino acids  
; FEATURE:  
; NAME/KEY: MISC FEATURE  
; LOCATION: (180)  
; OTHER INFORMATION: Xaa equals any of the twenty naturally occurring L-amino acids  
US-10-264-049-4308

Query Match 100.0%; Score 82; DB 15; Length 201;  
Best Local Similarity 100.0%; Pred. No. 9.7e-06;  
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GRPRHQGVVMGQK 15  
Db 53 GRPRHQGVVMGQK 67  
|||||

## RESULT 4

US-10-264-049-4148  
; Sequence 4148, Application US/10264049  
; Publication No. US20040005579A1  
; GENERAL INFORMATION:  
; APPLICANT: Birse et al.  
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies  
; FILE REFERENCE: PA133P1  
; CURRENT APPLICATION NUMBER: US/10/264,049  
; CURRENT FILING DATE: 2002-10-04  
; PRIOR APPLICATION NUMBER: PCT/US01/18569  
; PRIOR FILING DATE: 2001-06-07  
; PRIOR APPLICATION NUMBER: US 60/209,467  
; PRIOR FILING DATE: 2000-06-07  
; NUMBER OF SEQ ID NOS: 4360  
; SOFTWARE: PatentIn Ver. 3.1  
; SEQ ID NO 4148  
; LENGTH: 204  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-10-264-049-4148

Query Match 100.0%; Score 82; DB 15; Length 204;  
Best Local Similarity 100.0%; Pred. No. 9.9e-06;  
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GRPRHQGVVMGQK 15  
Db 56 GRPRHQGVVMGQK 70  
|||||

## RESULT 5

US-10-002-631C-56  
; Sequence 56, Application US/10002631C  
; Publication No. US20030157486A1  
; GENERAL INFORMATION:  
; APPLICANT: Graff, Jonathon M.  
; APPLICANT: Munster, Matthew  
; TITLE OF INVENTION: METHODS TO IDENTIFY SIGNAL SEQUENCES  
; FILE REFERENCE: A34943 090495.0243

Db	36	GRPRHGGVVMVGMGQK	50
<p>RESULT 8</p> <p>US-10-205-194-93</p> <p>Sequence 93, Application US/10205194</p> <p>Publication No. US20030134301A1</p> <p>GENERAL INFORMATION:</p> <p>APPLICANT: Warner-Lambert Company</p> <p>APPLICANT: Lee, Kevin</p> <p>APPLICANT: Dixon, Alistair</p> <p>APPLICANT: Brooksbank, Robert</p> <p>APPLICANT: Finnock, Robert</p> <p>TITLE OF INVENTION: Identification and Use of Molecules Implicated in Pain</p> <p>FILE REFERENCE: WL-A-018201</p> <p>CURRENT APPLICATION NUMBER: US/10/205,194</p> <p>CURRENT FILING DATE: 2000-07-24</p> <p>PRIOR APPLICATION NUMBER: GB 0118354.0</p> <p>PRIOR FILING DATE: 2001-07-27</p> <p>NUMBER OF SEQ ID NOS: 177</p> <p>SOFTWARE: PatentIn Ver. 2.1</p> <p>SEQ ID NO 93</p> <p>LENGTH: 375</p> <p>TYPE: PRT</p> <p>ORGANISM: Rattus rattus</p> <p>FEATURE:</p> <p>OTHER INFORMATION: Cytoplasmic gamma isoform of actin</p> <p>US-10-205-194-93</p>			
Query Match	100.0%	Score 82;	DB 14; Length 375;
Best Local Similarity	100.0%;	Pred. No. 1.8e-05;	
Matches	15;	Conservative 0;	Mismatches 0; Indels 0; Gaps 0;
QY	1	GRPRHGGVVMVGMGQK 15	
Db	36	GRPRHGGVVMVGMGQK 50	
<p>RESULT 9</p> <p>US-10-316-253-88</p> <p>Sequence 88, Application US/10316253</p> <p>Publication No. US20030162706A1</p> <p>GENERAL INFORMATION:</p> <p>APPLICANT: The Procter &amp; Gamble Company</p> <p>APPLICANT: Peters, Kevin</p> <p>APPLICANT: Thompson, Larry</p> <p>APPLICANT: Wang, Feng</p> <p>APPLICANT: Greis, Kenneth</p> <p>TITLE OF INVENTION: Angiogenesis Modulating Proteins</p> <p>FILE REFERENCE: 8865M</p> <p>CURRENT APPLICATION NUMBER: US/10/316,253</p> <p>CURRENT FILING DATE: 2002-12-10</p> <p>PRIOR APPLICATION NUMBER: US 60/355,295</p> <p>PRIOR FILING DATE: 2002-02-08</p> <p>NUMBER OF SEQ ID NOS: 308</p> <p>SOFTWARE: PatentIn version 3.1</p> <p>SEQ ID NO 88</p> <p>LENGTH: 375</p> <p>TYPE: PRT</p> <p>ORGANISM: Rattus norvegicus</p> <p>US-10-316-253-88</p>			
Query Match	100.0%	Score 82;	DB 14; Length 375;
Best Local Similarity	100.0%;	Pred. No. 1.8e-05;	
Matches	15;	Conservative 0;	Mismatches 0; Indels 0; Gaps 0;
QY	1	GRPRHGGVVMVGMGQK 15	
Db	36	GRPRHGGVVMVGMGQK 50	
<p>RESULT 10</p> <p>US-10-369-493-5927</p>			

; Sequence 5927, Application US/10369493  
; Publication No. US20030233675A1  
; GENERAL INFORMATION:  
; APPLICANT: Cao, Yongwei  
; APPLICANT: Hinkle, Gregory J.  
; APPLICANT: Slater, Steven C.  
; APPLICANT: Goldman, Barry S.  
; APPLICANT: Chen, Xiaofeng  
; TITLE OF INVENTION: EXPRESSION OF MICROBIAL PROTEINS IN PLANTS FOR PRODUCTION OF  
; TITLE OF INVENTION: PLANTS WITH IMPROVED PROPERTIES  
; FILE REFERENCE: 38-10(52052)B  
; CURRENT APPLICATION NUMBER: US/10/369,493  
; CURRENT FILING DATE: 2003-02-28  
; PRIOR APPLICATION NUMBER: US 60/360,039  
; PRIOR FILING DATE: 2002-02-21  
; NUMBER OF SEQ ID NOS: 47374  
; SEQ ID NO 5927  
; LENGTH: 375  
; TYPE: PRT  
; ORGANISM: Caenorhabditis elegans  
US-10-369-493-5927

Query Match 100.0%; Score 82; DB 15; Length 375;  
Best Local Similarity 100.0%; Pred. No. 1.8e-05;  
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
  
Qy 1 GRPRHQGVWVGQK 15  
Db 36 GRPRHQGVWVGQK 50

RESULT 11  
US-10-205-331-94  
; Sequence 94, Application US/10205331  
; Publication No. US20040058326A1  
; GENERAL INFORMATION:  
; APPLICANT: Lee, Kevin  
; APPLICANT: Dixon, Alistair  
; APPLICANT: Brooksbank, Robert  
; APPLICANT: Pinnock, Robert  
; TITLE OF INVENTION: Identification and Use of Molecules Implicated in Pain  
; FILE REFERENCE: WL-A-018199  
; CURRENT APPLICATION NUMBER: US/10/205,331  
; CURRENT FILING DATE: 2002-07-24  
; PRIOR APPLICATION NUMBER: GB 0118354.0  
; PRIOR FILING DATE: 2001-07-27  
; NUMBER OF SEQ ID NOS: 117  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 94  
; LENGTH: 375  
; TYPE: PRT  
; ORGANISM: Rattus norvegicus  
; FEATURE:  
; OTHER INFORMATION: Cytoplasmic beta-actin  
US-10-205-331-94

Query Match 100.0%; Score 82; DB 15; Length 375;  
Best Local Similarity 100.0%; Pred. No. 1.8e-05;  
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
  
Qy 1 GRPRHQGVWVGQK 15  
Db 36 GRPRHQGVWVGQK 50

RESULT 12  
US-10-260-708-82  
; Sequence 82, Application US/10260708  
; Publication No. US20040063101A1  
; GENERAL INFORMATION:  
; APPLICANT: Scanlan, Matthew  
; APPLICANT: Lee, Sang-Yull

; APPLICANT: Old, Lloyd  
; TITLE OF INVENTION: Human Sarcoma-Associated Antigens  
; FILE REFERENCE: L00461/70138  
; CURRENT APPLICATION NUMBER: US/10/260,708  
; CURRENT FILING DATE: 2002-09-30  
; NUMBER OF SEQ ID NOS: 96  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 82  
; LENGTH: 375  
; TYPE: PRT  
; ORGANISM: homo sapiens  
US-10-260-708-82

Query Match 100.0%; Score 82; DB 15; Length 375;  
Best Local Similarity 100.0%; Pred. No. 1.8e-05;  
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
  
Qy 1 GRPRHQGVWVGQK 15  
Db 36 GRPRHQGVWVGQK 50

RESULT 13  
US-10-341-434-101  
; Sequence 101, Application US/10341434  
; Publication No. US20030215835A1  
; GENERAL INFORMATION:  
; APPLICANT: Origene Technologies  
; TITLE OF INVENTION: Differentially Regulated Prostate Cancer Genes  
; FILE REFERENCE: 9U 204 205 R1  
; CURRENT APPLICATION NUMBER: US/10/341,434  
; CURRENT FILING DATE: 2003-07-18  
; PRIOR APPLICATION NUMBER: US 60/348,164  
; PRIOR FILING DATE: 2002-01-15  
; PRIOR APPLICATION NUMBER: US 60/348,119  
; PRIOR FILING DATE: 2002-01-15  
; NUMBER OF SEQ ID NOS: 238  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 101  
; LENGTH: 376  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-10-341-434-101

Query Match 100.0%; Score 82; DB 15; Length 376;  
Best Local Similarity 100.0%; Pred. No. 1.8e-05;  
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
  
Qy 1 GRPRHQGVWVGQK 15  
Db 37 GRPRHQGVWVGQK 51

RESULT 14  
US-09-919-172-33  
; Sequence 33, Application US/09919172  
; Patent No. US20020119463A1  
; GENERAL INFORMATION:  
; APPLICANT: Faris, Mary  
; APPLICANT: Turner, Christopher M.  
; TITLE OF INVENTION: PROSTATE CANCER MARKERS  
; FILE REFERENCE: PA-0036 US  
; CURRENT APPLICATION NUMBER: US/09/919,172  
; CURRENT FILING DATE: 2001-07-30  
; PRIOR APPLICATION NUMBER: 60/222,469  
; PRIOR FILING DATE: 2000-07-28  
; NUMBER OF SEQ ID NOS: 102  
; SOFTWARE: PERL Program  
; SEQ ID NO 33  
; LENGTH: 377  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
; FEATURE:

NAME/KEY: misc feature  
OTHER INFORMATION: Incyte ID No. US20020119463A1 1709118CD1  
US-09-19-172-33

Query Match 100.0%; Score 82; DB 9; Length 377;  
Best Local Similarity 100.0%; Pred. No. 1.9e-05;  
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GRPRHQGVVGMGQK 15  
| | | | | | | | | | | | | | |  
Db 38 GRPRHQGVVGMGQK 52

RESULT 15  
US-10-236-031B-64  
Sequence 64, Application US/10236031B  
Publication No. US20030219760A1  
GENERAL INFORMATION:  
APPLICANT: Gordon, Gavin J.  
APPLICANT: Jensen, Roderick V.  
APPLICANT: Gullans, Steven R.  
APPLICANT: Bueno, Raphael  
TITLE OF INVENTION: Diagnostic and Prognostic Tests  
FILE REFERENCE: B00801/70265 (JRV/JAV)  
CURRENT APPLICATION NUMBER: US/10/236,031B  
CURRENT FILING DATE: 2002-09-05  
PRIOR APPLICATION NUMBER: US 60/317,389  
PRIOR FILING DATE: 2001-09-05  
PRIOR APPLICATION NUMBER: US 60/407,431  
PRIOR FILING DATE: 2002-08-30  
NUMBER OF SEQ ID NOS: 102  
SOFTWARE: PatentIn version 3.1  
SEQ ID NO 64  
LENGTH: 377  
TYPE: PRT  
ORGANISM: Homo sapiens  
US-10-236-031B-64

Query Match 100.0%; Score 82; DB 15; Length 377;  
Best Local Similarity 100.0%; Pred. No. 1.9e-05;  
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GRPRHQGVVGMGQK 15  
| | | | | | | | | | | | | | |  
Db 38 GRPRHQGVVGMGQK 52

Search completed: April 8, 2005, 12:50:57  
Job time : 42.6667 secs

**THIS PAGE BLANK (USPTO)**

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: April 8, 2005, 09:58:56 ; Search time 9.33333 Seconds  
(without alignments)  
154.634 Million cell updates/sec

Title: US-09-423-351C-3  
Perfect score: 82  
Sequence: 1 GRPRHQGVWVGWGQK 15

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 283416 seqs, 96216763 residues  
Total number of hits satisfying chosen parameters: 283416

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000  
Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : PIR 79:\*  
1: pir1.\*  
2: pir2.\*  
3: pir3.\*  
4: pir4.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	82	100.0	86	C43616	actin beta, cytosol
2	82	100.0	137	A28258	actin 5C - fruit f
3	82	100.0	158	I49465	alpha-cardiac acti
4	82	100.0	213	A61043	actin CAL5 - sea s
5	82	100.0	308	A03000	actin 3 - fruit fl
6	82	100.0	328	S05430	actin beta - grass
7	82	100.0	349	B25819	actin, fetal skele
8	82	100.0	362	A26559	actin type 5, cyto
9	82	100.0	374	ATBOB	actin beta - bovin
10	82	100.0	374	ATBOG	actin gamma - bovi
11	82	100.0	374	JCS818	gamma-actin - huma
12	82	100.0	375	ATBOSM	actin, aortic smoo
13	82	100.0	375	ATRB	actin, skeletal mu
14	82	100.0	375	ATRTC	actin beta - rat
15	82	100.0	375	A48324	actin beta, cytosk
16	82	100.0	375	ATCHB	actin beta - chick
17	82	100.0	375	ATHUB	actin beta - human
18	82	100.0	375	ATHUG	actin gamma 1 - hu
19	82	100.0	375	ATMSG	actin beta - mouse
20	82	100.0	375	ATRSB	actin gamma - mous
21	82	100.0	375	ATRB2	actin beta, non-mu
22	82	100.0	375	S11222	actin gamma, cytos
23	82	100.0	375	T25272	hypothetical prote
24	82	100.0	375	S71125	actin beta-2, cyto
25	82	100.0	375	S71124	actin beta-1, cyto
26	82	100.0	375	S71126	actin beta, cytosol
27	82	100.0	375	A55001	actin beta - goose
28	82	100.0	375	A54728	actin alpha, cardi
29	82	100.0	375	S70377	actin - Phaffia rh

30	82	100.0	376	1	ATFF7	actin 7 - fruit fl
31	82	100.0	376	1	A43552	actin gamma, cytos
32	82	100.0	376	1	ATCHSM	actin gamma, smoot
33	82	100.0	376	1	ATFF8	actin 8 - fruit fl
34	82	100.0	376	1	ATURS	actin Cy1 - sea ur
35	82	100.0	376	2	A48449	Actin-1A - nematoe
36	82	100.0	376	2	S04538	actin 87E - fruit
37	82	100.0	376	2	JC1246	actin - fruit fly
38	82	100.0	376	2	JS0189	actin, cytosolic -
39	82	100.0	376	2	JS0190	actin, muscle - st
40	82	100.0	376	2	S07288	actin 15A - sea ur
41	82	100.0	376	2	S09578	actin - sea urchin
42	82	100.0	376	2	A40261	actin gamma, enter
43	82	100.0	376	2	JQ0154	actin - Hydra atte
44	82	100.0	376	2	JN0832	actin (clone gen3)
45	82	100.0	376	2	JN0833	actin (clones Ia a

## ALIGNMENTS

### RESULT 1

C43616  
actin beta, cytosolic - chicken (fragment)  
C/Species: Gallus gallus (Chicken)  
C/Date: 11-Dec-1992 #sequence\_revision 11-Dec-1992 #text\_change 09-Jul-2004  
C/Accession: C43616  
R/Paterson, B.M.; Eldridge, J.D.  
Science 224, 1436-1438, 1984  
A/Title: alpha-Cardiac actin is the major sarcomeric isoform expressed in embryonic avia  
A/Reference number: A43616; MUID:84223949; PMID:6729461  
A/Accession: C43616  
A/Status: preliminary  
A/Molecule type: DNA  
A/Residues: 1-86 <PAR>  
A/Cross-references: UNIPROT:Q90736; GB:K02173; NID:g211054; PIDN:AAA98513.1; PID:g211055  
C/Superfamily: actin  
C/Keywords: cytosol; methylated amino acid  
F,7/3/Modified site: 3'-methylhistidine (His) #status predicted

Query Match 100.0%; Score 82; DB 2; Length 86;  
Best Local Similarity 100.0%; Pred. No. 7e-07;  
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GRPRHQGVWVGWGQK 15  
DB 36 GRPRHQGVWVGWGQK 50

### RESULT 2

A28258  
actin 5C - fruit fly (Drosophila melanogaster)  
C/Species: Drosophila melanogaster  
C/Date: 30-Sep-1989 #sequence\_revision 30-Sep-1989 #text\_change 09-Jul-2004  
C/Accession: A28258  
R/Vigoreaux, J.O.; Tobin, S.L.  
Genes Dev. 1, 1161-1171, 1987  
A/Title: Stage-specific selection of alternative transcriptional initiation sites from  
A/Reference number: A28258; MUID:88112795; PMID:3123314  
A/Accession: A28258  
A/Molecule type: mRNA  
A/Residues: 1-137 <VIG>  
A/Cross-references: UNIPROT:P10987  
A/Note: the authors translated the codon GAG for residue 96 as Gly  
C/Genetics:  
A/Gene: FlyBase:Act5C  
A/Cross-references: FlyBase:FBgn0000042  
C/Superfamily: actin  
C/Keywords: methylated amino acid  
F,7/4/Modified site: 3'-methylhistidine (His) #status predicted

Query Match 100.0%; Score 82; DB 2; Length 137;  
Best Local Similarity 100.0%; Pred. No. 1.1e-06;

Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GRPRHQGVWVGQK 15  
 |||||  
 Db 37 GRPRHQGVWVGQK 51

RESULT 3  
 I49465  
 A:Title: alpha-cardiac actin - mouse (fragment)  
 C:Species: Mus musculus (house mouse)  
 C>Date: 02-Jul-1996 #sequence\_revision 02-Jul-1996 #text\_change 09-Jul-2004  
 C:Accession: I49465; I49466  
 R:Garner, I.; Minty, A.J.; Alonso, S.; Barton, P.J.R.; Buckingham, M.E.  
 EMBL J. 5, 2559-2567, 1986  
 A:Title: A 5' duplication of the alpha-cardiac actin gene in BALB/c mice is associated w  
 A:Reference number: I49465; MUID:87053822; PMID:3023046  
 A:Accession: I49465  
 A:Status: translated from GB/EMBL/DBJ  
 A:Molecule type: DNA  
 A:Gene: FlyBase:Act42A  
 A:Residues: 1-158 <RES>  
 A:CROSS-references: UNIPROT:Q61273; GB:M26775; NID:gl91646; PIDN:AAA37165.1; PID:g553858  
 A:Accession: I49466  
 A:Status: translated from GB/EMBL/DBJ  
 A:Molecule type: DNA  
 A:Residues: 8-158 <RE2>  
 A:CROSS-references: GB:M26776; NID:gl91649; PIDN:AAA37166.1; PID:g553859  
 A:Experimental source: adult cardiac muscle, BALB/c mice  
 C:Genetics:  
 A:Introns: 50/3  
 C:Superfamily: actin  
 C:Keywords: cardiac muscle; heart

Query Match 100.0%; Score 82; DB 2; Length 158;  
 Best Local Similarity 100.0%; Pred. No. 1.3e-06;  
 Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GRPRHQGVWVGQK 15  
 |||||  
 Db 45 GRPRHQGVWVGQK 59

RESULT 4  
 A61043  
 A:Title: actin CAL5 - sea squirt (Styela clava) (fragments)  
 C:Species: Styela clava  
 C>Date: 31-Dec-1993 #sequence\_revision 31-Dec-1993 #text\_change 09-Jul-2004  
 C:Accession: A61043  
 R:Beach, R.L.; Jeffery, W.R.  
 Dev. Genet. 11, 2-14, 1990  
 A:Title: Temporal and spatial expression of a cytoskeletal actin gene in the ascidian St  
 A:Reference number: A61043; MUID:90298580; PMID:2361333  
 A:Accession: A61043  
 A:Status: not compared with conceptual translation  
 A:Molecule type: mRNA  
 A:Residues: 1-213 <BEA>  
 A:CROSS-references: UNIPROT:Q7M3Y7  
 C:Comment: This sequence is expressed in cells undergoing rapid cell division.  
 C:Superfamily: actin  
 C:Keywords: cytoskeleton; methylated amino acid; mitosis; structural protein  
 F:73/Modified site: 3'-methylhistidine (His) #status predicted

Query Match 100.0%; Score 82; DB 2; Length 213;  
 Best Local Similarity 100.0%; Pred. No. 1.7e-06;  
 Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GRPRHQGVWVGQK 15  
 |||||  
 Db 36 GRPRHQGVWVGQK 50

RESULT 5  
 A03000

actin 3 - fruit fly (Drosophila melanogaster) (fragments)  
 C:Species: Drosophila melanogaster  
 C>Date: 31-Mar-1991 #sequence\_revision 31-Mar-1991 #text\_change 09-Jul-2004  
 C:Accession: A03000  
 R:Fyrberg, E.A.; Bond, B.J.; Hershey, N.D.; Mixter, K.S.; Davidson, N.  
 Cell 24, 107-116, 1981  
 A:Title: The actin genes of Drosophila: protein coding regions are highly conserved but  
 A:Reference number: A03000; MUID:81210174; PMID:6263481  
 A:Accession: A03000  
 A:Molecule type: DNA  
 A:Residues: 1-308 <FYR>  
 A:CROSS-references: UNIPROT:P02572  
 A:Note: there are 68 unsequenced residues between positions 160 and 161. Partial sequence  
 A:Note: the authors translated the codon GTT for residue 263 as Ile  
 C:Genetics:  
 A:Gene: FlyBase:Act42A  
 A:CROSS-references: FlyBase:FBgn0000043  
 A:Map position: 42A  
 C:Superfamily: actin  
 C:Keywords: methylated amino acid  
 F:74/Modified site: 3'-methylhistidine (His) #status predicted

Query Match 100.0%; Score 82; DB 2; Length 308;  
 Best Local Similarity 100.0%; Pred. No. 2.5e-06;  
 Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GRPRHQGVWVGQK 15  
 |||||  
 Db 37 GRPRHQGVWVGQK 51

RESULT 6  
 S05430  
 A:Title: actin beta - grass carp  
 C:Species: Ctenopharyngodon idella (grass carp)  
 C>Date: 07-Jun-1990 #sequence\_revision 07-Jun-1990 #text\_change 09-Jul-2004  
 C:Accession: S05430  
 R:Liu, Z.; Zhu, Z.; Roberg, K.; Faras, A.J.; Guise, K.S.; Kapuscinski, A.R.; Hackett, P.  
 Nucleic Acids Res. 17, 5850, 1989  
 A:Title: The beta-actin gene of carp (Ctenopharyngodon idella).  
 A:Reference number: S05430; MUID:89345185; PMID:2762162  
 A:Accession: S05430  
 A:Status: translation not shown  
 A:Molecule type: DNA  
 A:Residues: 1-328 <LIU>  
 A:CROSS-references: UNIPROT:P83751; EMBL:M25013  
 C:Genetics:  
 A:Introns: 41/3; 121/3; 268/1  
 C:Superfamily: actin  
 C:Keywords: cytoskeleton; methylated amino acid  
 F:73/Modified site: 3'-methylhistidine (His) #status predicted

Query Match 100.0%; Score 82; DB 2; Length 328;  
 Best Local Similarity 100.0%; Pred. No. 2.7e-06;  
 Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GRPRHQGVWVGQK 15  
 |||||  
 Db 36 GRPRHQGVWVGQK 50

RESULT 7  
 B25819  
 A:Title: actin, fetal skeletal/adult cardiac muscle - mouse (fragment)  
 C:Species: Mus musculus (house mouse)  
 C>Date: 23-Aug-1987 #sequence\_revision 23-Aug-1987 #text\_change 09-Jul-2004  
 C:Accession: B25819  
 R:Alonso, S.; Minty, A.; Bourlet, Y.; Buckingham, M.  
 J. Mol. Evol. 23, 11-22, 1986  
 A:Title: Comparison of three actin-coding sequences in the mouse; evolutionary relations  
 A:Reference number: A25819; MUID:86200234; PMID:3084797  
 A:Accession: B25819  
 A:Molecule type: mRNA



A:Residues: 1-349 <ALO>  
A:Cross-references: UNIPROT:061275; GB:X03767; GB:J00381; GB:M10652; NID:G49869; PIDN:CA  
C:Superfamily: actin  
C:Keywords: cardiac muscle; heart; methylated amino acid; muscle; skeletal muscle  
F:47/Modified site: 3'-methylhistidine (His) #status predicted

Query Match 100.0%; Score 82; DB 2; Length 349;  
Best Local Similarity 100.0%; Pred. No. 2.8e-06;  
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GRPRHGVVGVGMGQK 15  
|||||  
Db 10 GRPRHGVVGVGMGQK 24  
|||||

RESULT 8  
A26559  
actin type 5, cytosolic - chicken  
C:Species: Gallus gallus (chicken)  
C:Date: 09-Sep-1987 #sequence\_revision 09-Sep-1987 #text\_change 05-Dec-1997  
C:Accession: A26559  
R:Bergsma, D.J.; Chang, K.S.; Schwartz, R.J.  
Mol. Cell. Biol. 5, 1151-1162, 1985  
A:Reference number: A26559; MUID:85213487; PMID:4000121  
A:Accession: A26559  
A:Molecule type: DNA  
A:Residues: 1-362 <BER>  
C:Superfamily: actin  
C:Keywords: cytosol; methylated amino acid  
F:74/Modified site: 3'-methylhistidine (His) #status predicted

Query Match 100.0%; Score 82; DB 2; Length 362;  
Best Local Similarity 100.0%; Pred. No. 3e-06;  
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GRPRHGVVGVGMGQK 15  
|||||  
Db 37 GRPRHGVVGVGMGQK 51  
|||||

RESULT 9  
ATBOG  
actin beta - bovine (tentative sequence)  
C:Species: Bos primigenius taurus (cattle)  
C:Date: 30-Sep-1991 #sequence\_revision 30-Sep-1991 #text\_change 09-Jul-2004  
C:Accession: E14185; A39105; A02999; A14185  
R:Vandekerckhove, J.; Weber, K.  
Eur. J. Biochem. 90, 451-462, 1978  
A:Title: Actin amino-acid sequences. Comparison of actins from calf thymus, bovine brain  
A:Reference number: A14185; MUID:79045349; PMID:213279  
A:Accession: E14185  
A:Molecule type: protein  
A:Residues: 1-374 <VAN>  
A:Cross-references: UNIPROT:P60712  
A:Note: Only peptides that differed in composition from the corresponding peptides of ra  
R:Degen, J.L.; Neubauer, M.G.; Degen, S.J.F.; Seyfried, C.E.; Morris, D.R.  
J. Biol. Chem. 258, 12153-12162, 1983  
A:Title: Regulation of protein synthesis in mitogen-activated bovine lymphocytes. Analysis  
A:Reference number: A39105; MUID:84032385; PMID:6195151  
A:Accession: A39105  
A:Molecule type: mRNA  
A:Residues: 76-227;344-374 <DEG>  
A:Cross-references: GB:K00622; GB:K00623  
A:Note: actins beta and gamma were not distinguished in this study  
C:Comment: Vertebrate nonmuscle cells contain two highly conserved cytoskeletal actins,  
C:Superfamily: actin  
F:1/Modified site: blocked amino end (Asp) (probably acetylated) #status experimental  
F:72/Modified site: 3'-methylhistidine (His) #status predicted

Query Match 100.0%; Score 82; DB 1; Length 374;  
Best Local Similarity 100.0%; Pred. No. 3.1e-06;  
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GRPRHGVVGVGMGQK 15  
|||||  
Db 35 GRPRHGVVGVGMGQK 49  
|||||

RESULT 10  
ATBOG  
actin gamma - bovine (tentative sequence)  
C:Species: Bos primigenius taurus (cattle)  
C:Date: 30-Sep-1991 #sequence\_revision 30-Sep-1991 #text\_change 09-Jul-2004  
C:Accession: E14185; A02999  
R:Vandekerckhove, J.; Weber, K.  
Eur. J. Biochem. 90, 451-462, 1978  
A:Title: Actin amino-acid sequences. Comparison of actins from calf thymus, bovine brain  
A:Reference number: A14185; MUID:79045349; PMID:213279  
A:Accession: E14185  
A:Molecule type: protein  
A:Residues: 1-374 <VAN>  
A:Cross-references: UNIPROT:P02571  
A:Note: Only peptides that differed in composition from the corresponding peptides of ra  
C:Comment: Vertebrate nonmuscle cells contain two highly conserved cytoskeletal actins,  
C:Superfamily: actin  
C:Keywords: blocked amino end; cell motility; cytoskeleton; methylated amino acid; micro  
F:1/Modified site: blocked amino end (Glu) (probably acetylated) #status experimental  
F:72/Modified site: 3'-methylhistidine (His) #status predicted

Query Match 100.0%; Score 82; DB 1; Length 374;  
Best Local Similarity 100.0%; Pred. No. 3.1e-06;  
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GRPRHGVVGVGMGQK 15  
|||||  
Db 35 GRPRHGVVGVGMGQK 49  
|||||

RESULT 11  
JC5818  
gamma-actin - human  
C:Species: Homo sapiens (man)  
C:Date: 20-Feb-1998 #sequence\_revision 13-Mar-1998 #text\_change 09-Jul-2004  
C:Accession: JC5818; PC4501  
R:Hauschildt, S.; Schwarz, C.; Heine, H.; Ulmer, A.J.; Flad, H.D.; Rietschel, E.T.; Jens  
Biochem. Biophys. Res. Commun. 241, 670-674, 1997  
A:Title: Actin: A target of lipopolysaccharide-induced phosphorylation in human monocytes  
A:Reference number: JC5818; MUID:98096379; PMID:9434766  
A:Accession: JC5818  
A:Molecule type: protein  
A:Residues: 1-374 <HAU>  
A:Experimental source: monocyte  
A:Cross-references: UNIPROT:P02571  
A:Accession: PC4501  
A:Molecule type: protein  
A:Residues: 1-61;84-112;147-190;196-209;215-253;335-358 <HA2>  
A:Experimental source: monocyte  
C:Comment: This protein is involved in a signal transduction that eventually leads to mo  
C:Superfamily: actin

Query Match 100.0%; Score 82; DB 2; Length 374;  
Best Local Similarity 100.0%; Pred. No. 3.1e-06;  
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GRPRHGVVGVGMGQK 15  
|||||  
Db 35 GRPRHGVVGVGMGQK 49  
|||||

RESULT 12  
ATBOG  
actin, aortic smooth muscle - bovine  
C:Species: Bos primigenius taurus (cattle)  
C:Date: 30-Jun-1987 #sequence\_revision 30-Jun-1987 #text\_change 06-Sep-1996  
C:Accession: A02997; S13480

A;Residues: 48-54;68-72;235-243 <STR>  
A;Experimental source: skeletal muscle  
R;Bertrand, R.; Derancourt, J.; Kassab, R.  
FEBS Lett. 345, 113-119, 1994  
A;Title: The covalent maleinidobenzoyl-actin-myosin head complex. Cross-linking of the 5  
A;Reference number: S44393; MUID:94259162; PMID:8200441  
A;Accession: S44393  
A;Molecule type: protein  
A;Residues: 48-64 <BER>  
A;Experimental source: skeletal muscle  
C;Superfamily: actin  
C;Keywords: acetylated amino end; ATP binding; methylated amino acid; muscle contraction  
F;1-375/Product: actin #status experimental <MAT>  
F;1/Modified site: acetylated amino end (Asp) #status experimental  
F;73/Modified site: 3'-methylhistidine (His) #status experimental

Query Match 100.0%; Score 82; DB 1; Length 375;  
Best Local Similarity 100.0%; Pred. No. 3.1e-06;  
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GRPRHQGVVGMGQK 15  
|||  
Db 36 GRPRHQGVVGMGQK 50

RESULT 14

ATRTC  
actin beta - rat  
C;Species: Rattus norvegicus (Norway rat)  
C;Date: 13-Jun-1983 #sequence\_revision 13-Jun-1983 #text\_change 22-Jun-1999  
C;Accession: A38571; A02999  
R;Nudel, U.; Zakut, R.; Shani, M.; Neuman, S.; Levy, Z.; Yaffe, D.  
Nucleic Acids Res. 11, 1759-1771, 1983  
A;Title: The nucleotide sequence of the rat cytoplasmic beta-actin gene.  
A;Reference number: A38571; MUID:83168920; PMID:6300777  
A;Accession: A38571  
A;Molecule type: DNA  
A;Residues: 1-375 <NUD>  
A;Cross-references: GB:J00691; NID:G202653; PIDN:AAA40657.1; PID:G202654  
C;Comment: Vertebrate nonmuscle cells contain two highly conserved cytoskeletal actins,  $\alpha$   
C;Genetics: 41/3; 121/3; 268/1; 328/3  
C;Introns: 41/3; 121/3; 268/1; 328/3  
C;Superfamily: actin  
C;Keywords: cell motility; cytoskeleton; methylated amino acid; microfilament; mitosis;  
F;2-375/Product: actin beta #status predicted <MAT>  
F;73/Modified site: 3'-methylhistidine (His) #status predicted

Query Match 100.0%; Score 82; DB 1; Length 375;  
Best Local Similarity 100.0%; Pred. No. 3.1e-06;  
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GRPRHQGVVGMGQK 15  
|||  
Db 36 GRPRHQGVVGMGQK 50

RESULT 15

actin beta, cytoskeletal - common carp  
C;Species: Cyprinus carpio (common carp)  
C;Date: 03-Feb-1994 #sequence\_revision 11-Apr-1997 #text\_change 09-Jul-2004  
C;Accession: A48324  
R;Liu, Z.; Zhu, Z.; Roberg, K.; Faras, A.; Guise, K.; Kapuscinski, A.R.; Hackett, P.B.  
DNA Seq. 1, 125-136, 1990  
A;Title: Isolation and characterization of beta-actin gene of carp (Cyprinus carpio).  
A;Reference number: A48324; MUID:92190540; PMID:2134183  
A;Accession: A48324  
A;Molecule type: DNA  
A;Residues: 1-375 <LIU>  
A;Cross-references: UNIPROT:P83750; GB:M24113; NID:G213041; PIDN:AAA68886.1; PID:G213042  
A;Note: the authors translated the codon TTC for residue 21 as Pro, AAG for residue 50  
7 as Pro  
A;Note: the authors failed to translated the codon GGT for residue 42 as Gly

C;Comment: Vertebrate nonmuscle cells contain two highly conserved cytoskeletal actins,

**C;Genetics:**

A; Introns: 41/3; 121/3; 268/3; 328/3

C;Superfamily: actin

C;Keywords: acetylated amino end; cell motility; cytoskeleton; methylated amino acid; microfilament; superparamagnetic contrast agent

E;2-375/Product: actin beta, cytoskeletal #status predicted <MAT>

F;2/Modified site: acetylated amino end (Asp) (in mature form) #status predicted

P;73/Modified site: 3'-methylhistidine (His) #status predicted

Query Match 100.0%; Score 82; DB 1; Length 375;

Query Match: 100.0%; Score 82; DS 1; Length 379;  
Best Local Similarity 100.0%; Pred. No. 3.1e-06;

```

DESC LOCAL Similarity 100.0%, FREQ. NO. 3.1E+00;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

Qy 1 GRPRHQGMVGMGQK 15

QY  
T GRFHQGVIVGMBQR

Db 36 GRPRHOGVMVGMGOK

DD 36 GRPRHQGVMMVGMGQK 50

Search completed: April 8, 2005, 10:53:05

search completed: APR11  
Job time : 10.3333 secsv

**THIS PAGE BLANK (USPTO)**

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: April 8, 2005, 09:58:56 ; Search time 57.8667 Seconds  
(without alignments)  
132.739 Million cell updates/sec

Title: US-09-423-351C-3

Perfect score: 82  
Sequence: 1 GRPRHQGVVGMGQK 15

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 1612378 seqs, 512079187 residues

Total number of hits satisfying chosen parameters: 1612378

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : UniProt\_03.\*

1: uniprot\_sprot.\*

2: uniprot\_trembl.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	82	100.0	26	2 Q6Q298	Q6Q298 ovix aries
2	82	100.0	53	2 Q6AW41	Q6aw41 bombyx mori
3	82	100.0	78	2 Q61375	Q61375 schmidtea p
4	82	100.0	78	2 Q61376	Q61376 schmidtea p
5	82	100.0	78	2 Q61377	Q61377 schmidtea p
6	82	100.0	79	2 Q95L83	/ Q95L83 oryctolagus
7	82	100.0	80	2 Q9TWT4	Q9ttw4 bos taurus
8	82	100.0	86	2 Q90736	Q90736 gallus gall
9	82	100.0	91	2 Q72Z11	Q7zzil salvelinus
10	82	100.0	96	2 Q86CV4	Q86cv4 drosophila
11	82	100.0	96	2 Q86FV7	Q86fv7 drosophila
12	82	100.0	99	2 Q9QZB4	Q9qzb4 cavia porce
13	82	100.0	104	2 Q86SB7	Q86sb7 gryllus bim
14	82	100.0	108	2 Q18548	Q18548 lytechinus
15	82	100.0	108	2 Q18550	Q18550 lytechinus
16	82	100.0	117	2 Q726P7	Q726p7 elysia chlo
17	82	100.0	121	2 Q8JG71	Q8jg71 oreochromis
18	82	100.0	121	2 Q78CQ1	Q78cq1 pseudopleur
19	82	100.0	129	2 Q6VBE4	Q6vbe4 locusta mig
20	82	100.0	130	2 Q93298	Q93298 pseudopleur
21	82	100.0	139	2 Q9DFK2	Q9dfk2 gillichthys
22	82	100.0	144	2 Q90YX9	Q90yx9 fundulus he
23	82	100.0	150	2 Q804Y9	Q804y9 ictalurus p
24	82	100.0	151	2 Q61274	Q61274 mus musculu
25	82	100.0	151	2 Q9PSV5	Q9psv5 oryzias lat
26	82	100.0	158	2 Q61273	Q61273 mus musculu
27	82	100.0	164	1 ACT SPOLI	Q11212 spodoptera
28	82	100.0	165	2 Q17477	Q17477 halictis ru
29	82	100.0	181	2 Q8JG41	Q8jg41 dicentrarch
30	82	100.0	181	2 Q8AWC8	Q8awc8 chionodraco
31	82	100.0	181	2 Q8AWZ3	Q8awz3 argyropelec

32	82	100.0	181	2 Q9DEU3	Q9deu3 scophthalmu
33	82	100.0	184	2 Q11209	Q11209 canis famil
34	82	100.0	186	2 Q8X190	Q8x190 paxillus in
35	82	100.0	189	2 Q8SPX4	Q8spx4 canis famil
36	82	100.0	191	2 Q84G12	Q84g12 oxyuranus s
37	82	100.0	198	2 Q8Y250	Q8y250 pagrus majo
38	82	100.0	202	2 Q810R5	Q810r5 mus musculu
39	82	100.0	212	2 Q6ZYL2	Q6zyl2 arion lusit
40	82	100.0	213	2 Q7M3Y7	Q7m3y7 styela clav
41	82	100.0	225	2 Q64G13	Q64g13 oxyuranus s
42	82	100.0	235	2 Q7XB23	Q7xb23 lotharella
43	82	100.0	235	2 Q7XB24	Q7xb24 lotharella
44	82	100.0	235	2 Q7XB25	Q7xb25 lotharella
45	82	100.0	236	2 Q6RXK3	Q6rxk3 penaeus van

## ALIGNMENTS

### RESULT 1

Q6Q298 PRELIMINARY; PRT; 26 AA.  
AC Q6Q298;  
DT 05-JUL-2004 (Tremblrel. 27, Created)  
DT 05-JUL-2004 (Tremblrel. 27, Last sequence update)  
DT 05-JUL-2004 (Tremblrel. 27, Last annotation update)  
DE Beta actin (Fragment).  
OS Ovis aries (Sheep).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;  
OC Caprinae; Ovis.  
OX NCBI\_TaxID=9940;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC TISSUE=Abomasum;  
RA Shanmuganathan T., Hickford J.G.H., Savill M., Sykes A.R.;  
RL Submitted (MAR-2004) to the EMBL/GenBank/DBJ databases.  
DR EMBL; AY566300; AAS68014.1; -.  
FT NON\_TER 1  
FT NON\_TER 26  
SQ SEQUENCE 26 AA; 2802 MW; C75DA20C92E36C4B CRC64;

Query Match 100.0%; Score 82; DB 2; Length 26;  
Best Local Similarity 100.0%; Pred. No. 1.4e-06;  
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GRPRHQGVVGMGQK 15  
Db 7 GRPRHQGVVGMGQK 21

### RESULT 2

Q6AW41 PRELIMINARY; PRT; 53 AA.  
ID Q6AW41;  
AC Q6AW41;  
DT 25-OCT-2004 (Tremblrel. 28, Created)  
DT 25-OCT-2004 (Tremblrel. 28, Last sequence update)  
DT 25-OCT-2004 (Tremblrel. 28, Last annotation update)  
DE Cytoplasmic actin (Fragment).  
OS Bombyx mori (Silk moth).  
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;  
OC Neoptera; Endopterygota; Lepidoptera; Glossata; Ditrysia; Bombycoidea;  
OC Bombycidae; Bombyx.  
OX NCBI\_TaxID=7091;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=Daizo; TISSUE=silk gland;  
RA Shiomi K., Kajiuura Z., Nakagaki M., Yamashita O.;  
RT "Baculovirus-mediated efficient gene transfer into the central nervous system of the silkworm, Bombyx mori.";  
RL Nihon Sanshigaku Zasshi 72:149-155(2003).  
RN [2]  
RP SEQUENCE FROM N.A.

```

RC STRAIN=Daizo; TISSUE=Silk gland;
RA Shiomi K.;
RL Submitted (AUG-2004) to the EMBL/GenBank/DBJ databases.
DR EMBL; AB186491; BAD35130.1; -.
DR InterPro; IPR004000; Actin_like.
DR Pfam; PF00022; Actin; 1.
FT NON-TER 53
SQ SEQUENCE 53 AA; 5465 MW; 227AA8B4872EBA86 CRC64;

Query Match 100.0%; Score 82; DB 2; Length 53;
Best Local Similarity 100.0%; Pred. No. 2.8e-06;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GRPRHQGVWVGQK 15
Db 37 GRPRHQGVWVGQK 51

RESULT 3
O61375 PRELIMINARY; PRT; 78 AA.
AC O61375;
DT 01-AUG-1998 (TrEMBLrel. 07, Created)
DT 01-AUG-1998 (TrEMBLrel. 07, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE Actin 1 (Fragment).
GN Name=DpAct1;
OS Schmidtea polychroa.
OC Eukaryota; Metazoa; platyhelminthes; Turbellaria; Seriata; Tricladida;
OC Paludicola; Dugesidae; Schmidtea.
OX NCBI_TaxID=50054;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=98186847; PubMed=9518500; DOI=10.1093/nar/26.8.2031;
RA Fagotti A., Gabbiani G., Pascolini R., Neuville P.;
RT "Multiple isoform recovery (MIR)-PCR: a simple method for the
isolation of related mRNA isoforms.";
RL Nucleic Acids Res. 26:2031-2033(1998).
CC -!- FUNCTION: Actins are highly conserved proteins that are involved
in various types of cell motility and are ubiquitously expressed
in all eukaryotic cells (By similarity).
CC -!- SIMILARITY: Belongs to the actin family.
DR EMBL; AF027162; AAC38982.1; -.
DR HSP; P02577; INMI.
DR GO; GO:0005884; C:actin filament; IEA.
DR GO; GO:0003774; F:motor activity; IEA.
DR GO; GO:0005200; F:structural constituent of cytoskeleton; IEA.
DR InterPro; IPR004001; Actin.
DR Pfam; PF00022; Actin; 1.
DR PRINTS; PR00190; ACTIN.
DR PROSITE; PS00406; ACTINS_1; 1.
KW Structural protein.
FT NON-TER 78
SQ SEQUENCE 78 AA; 8205 MW; B9DFF9B28E4ADB0E CRC64;

Query Match 100.0%; Score 82; DB 2; Length 78;
Best Local Similarity 100.0%; Pred. No. 4.1e-06;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GRPRHQGVWVGQK 15
Db 37 GRPRHQGVWVGQK 51

RESULT 5
O61377 PRELIMINARY; PRT; 78 AA.
AC O61377;
DT 01-AUG-1998 (TrEMBLrel. 07, Created)
DT 01-AUG-1998 (TrEMBLrel. 07, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE Actin 3 (Fragment).
GN Name=DpAct3;
OS Schmidtea polychroa.
OC Eukaryota; Metazoa; platyhelminthes; Turbellaria; Seriata; Tricladida;
OC Paludicola; Dugesidae; Schmidtea.
OX NCBI_TaxID=50054;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=98186847; PubMed=9518500; DOI=10.1093/nar/26.8.2031;
RA Fagotti A., Gabbiani G., Pascolini R., Neuville P.;
RT "Multiple isoform recovery (MIR)-PCR: a simple method for the
isolation of related mRNA isoforms.";
RL Nucleic Acids Res. 26:2031-2033(1998).
CC -!- FUNCTION: Actins are highly conserved proteins that are involved
in various types of cell motility and are ubiquitously expressed
in all eukaryotic cells (By similarity).
CC -!- SIMILARITY: Belongs to the actin family.
DR EMBL; AF027161; AAC38981.1; -.
DR HSP; P02577; INMI.
DR GO; GO:0005884; C:actin filament; IEA.
DR GO; GO:0003774; F:motor activity; IEA.
DR GO; GO:0005200; F:structural constituent of cytoskeleton; IEA.
DR InterPro; IPR004001; Actin.
DR Pfam; PF00022; Actin; 1.
DR PRINTS; PR00190; ACTIN.
DR PROSITE; PS00406; ACTINS_1; 1.
KW Structural protein.
FT NON-TER 78
SQ SEQUENCE 78 AA; 8205 MW; 572BAB2698C24404 CRC64;

Query Match 100.0%; Score 82; DB 2; Length 78;
Best Local Similarity 100.0%; Pred. No. 4.1e-06;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GRPRHQGVWVGQK 15
Db 37 GRPRHQGVWVGQK 51

RESULT 4
O61376 PRELIMINARY; PRT; 78 AA.
AC O61376;
DT 01-AUG-1998 (TrEMBLrel. 07, Created)
DT 01-AUG-1998 (TrEMBLrel. 07, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE Actin 2 (Fragment).

```

```

GN Name=DpAct2;
OS Schmidtea polychroa.
OC Eukaryota; Metazoa; platyhelminthes; Turbellaria; Seriata; Tricladida;
OC Paludicola; Dugesidae; Schmidtea.
OX NCBI_TaxID=50054;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=98186847; PubMed=9518500; DOI=10.1093/nar/26.8.2031;
RA Fagotti A., Gabbiani G., Pascolini R., Neuville P.;
RT "Multiple isoform recovery (MIR)-PCR: a simple method for the
isolation of related mRNA isoforms.";
RL Nucleic Acids Res. 26:2031-2033(1998).
CC -!- FUNCTION: Actins are highly conserved proteins that are involved
in various types of cell motility and are ubiquitously expressed
in all eukaryotic cells (By similarity).
CC -!- SIMILARITY: Belongs to the actin family.
DR EMBL; AF027162; AAC38982.1; -.
DR HSP; P02577; INMI.
DR GO; GO:0005884; C:actin filament; IEA.
DR GO; GO:0003774; F:motor activity; IEA.
DR GO; GO:0005200; F:structural constituent of cytoskeleton; IEA.
DR InterPro; IPR004001; Actin.
DR Pfam; PF00022; Actin; 1.
DR PRINTS; PR00190; ACTIN.
DR PROSITE; PS00406; ACTINS_1; 1.
KW Structural protein.
FT NON-TER 78
SQ SEQUENCE 78 AA; 8265 MW; B9DFF9B28E4ADB0E CRC64;

Query Match 100.0%; Score 82; DB 2; Length 78;
Best Local Similarity 100.0%; Pred. No. 4.1e-06;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GRPRHQGVWVGQK 15
Db 37 GRPRHQGVWVGQK 51

RESULT 5
O61377 PRELIMINARY; PRT; 78 AA.
AC O61377;
DT 01-AUG-1998 (TrEMBLrel. 07, Created)
DT 01-AUG-1998 (TrEMBLrel. 07, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE Actin 3 (Fragment).
GN Name=DpAct3;
OS Schmidtea polychroa.
OC Eukaryota; Metazoa; platyhelminthes; Turbellaria; Seriata; Tricladida;
OC Paludicola; Dugesidae; Schmidtea.
OX NCBI_TaxID=50054;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=98186847; PubMed=9518500; DOI=10.1093/nar/26.8.2031;
RA Fagotti A., Gabbiani G., Pascolini R., Neuville P.;
RT "Multiple isoform recovery (MIR)-PCR: a simple method for the
isolation of related mRNA isoforms.";
RL Nucleic Acids Res. 26:2031-2033(1998).
CC -!- FUNCTION: Actins are highly conserved proteins that are involved
in various types of cell motility and are ubiquitously expressed
in all eukaryotic cells (By similarity).
CC -!- SIMILARITY: Belongs to the actin family.
DR EMBL; AF027163; AAC38983.1; -.
DR HSP; P02577; INMI.
DR GO; GO:0005884; C:actin filament; IEA.
DR GO; GO:0003774; F:motor activity; IEA.
DR GO; GO:0005200; F:structural constituent of cytoskeleton; IEA.
DR InterPro; IPR004001; Actin.
DR Pfam; PF00022; Actin; 1.
DR PRINTS; PR00190; ACTIN.
DR PROSITE; PS00406; ACTINS_1; 1.

```

```
KW Structural protein.
FT NON TER 78
SQ SEQUENCE 78 AA; 8299 MW; B9DFE1108E4ADBOE CRC64;

Query Match 100.0%; Score 82; DB 2; Length 78;
Best Local Similarity 100.0%; Pred. No. 4.1e-06;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GRPRHQGVVGMGQK 15
    |||||
Db 37 GRPRHQGVVGMGQK 51

RESULT 6
Q95L83
ID Q95L83 PRELIMINARY; PRT; 79 AA.
AC Q95L83;
DT 01-DEC-2001 (TrEMBLrel. 19, Created)
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE Beta-actin (Fragment)
OS Oryctolagus cuniculus (Rabbit)
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Lagomorpha; Leporidae; Oryctolagus.
OX NCBI_TaxID=9986;
RN [1]
RP SEQUENCE FROM N.A. PubMed=12388084;
RX MEDLINE=22306422; PubMed=12388084;
RA Wang D.-A., Du H., Jaggar J.H., Brindley D.N., Tigyi G.J.,
RA Watsky M.A.;
RT "Injury-elicited differential transcriptional regulation of
RT phospholipid growth factor receptors in the cornea.";
RL Am. J. Physiol. Cell Physiol. 283:C1646-C1654(2002).
CC -!- FUNCTION: Actins are highly conserved proteins that are involved
CC in various types of cell motility and are ubiquitously expressed
CC in all eukaryotic cells (By similarity).
CC -!- SIMILARITY: Belongs to the actin family.
DR EMBL; AF404278; AAL01885.1; -
DR HSSP; P02577; INMI.
DR GO; GO:0005884; C:actin filament; IEA.
DR GO; GO:0003774; F:motor activity; IEA.
DR GO; GO:0005200; F:structural constituent of cytoskeleton; IEA.
DR InterPro; IPR004001; Actin.
DR InterPro; IPR004000; Actin_like.
DR Pfam; PF00022; Actin; 1.
DR PRINTS; PR00190; ACTIN.
DR PROSITE; PS00406; ACTINS_1; 1.
KW Structural protein.
FT NON TER 1 79
SQ SEQUENCE 79 AA; 8420 MW; 05056E4E2AC092F8 CRC64;

Query Match 100.0%; Score 82; DB 2; Length 79;
Best Local Similarity 100.0%; Pred. No. 4.1e-06;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GRPRHQGVVGMGQK 15
    |||||
Db 38 GRPRHQGVVGMGQK 52

RESULT 7
Q9TTW4
ID Q9TTW4 PRELIMINARY; PRT; 80 AA.
AC Q9TTW4;
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE Beta actin (Fragment)
OS Bos taurus (Bovine)
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
OC Bovinae; Bos.
```

```
OX NCBI_TaxID=9913;
RN [1]
RP SEQUENCE FROM N.A.
RA John S.J., Bildeau-Goeseels S.;
RL Submitted (OCT-1999) to the EMBL/GenBank/DBJ databases.
CC -!- FUNCTION: Actins are highly conserved proteins that are involved
CC in various types of cell motility and are ubiquitously expressed
CC in all eukaryotic cells (By similarity).
CC -!- SIMILARITY: Belongs to the actin family.
DR EMBL; AF191490; AAF05984.1; -
DR HSSP; P02577; INMI.
DR GO; GO:0005884; C:actin filament; IEA.
DR GO; GO:0003774; F:motor activity; IEA.
DR GO; GO:0005200; F:structural constituent of cytoskeleton; IEA.
DR InterPro; IPR004001; Actin.
DR InterPro; IPR004000; Actin_like.
DR Pfam; PF00022; Actin; 1.
DR PRINTS; PR00190; ACTIN.
DR PROSITE; PS00406; ACTINS_1; 1.
KW Structural protein.
FT NON TER 1 80
SQ SEQUENCE 80 AA; 9133 MW; 47354ABB7010668D CRC64;

Query Match 100.0%; Score 82; DB 2; Length 80;
Best Local Similarity 100.0%; Pred. No. 4.2e-06;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GRPRHQGVVGMGQK 15
    |||||
Db 2 GRPRHQGVVGMGQK 16

RESULT 8
Q90736
ID Q90736 PRELIMINARY; PRT; 86 AA.
AC Q90736;
DT 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE Beta-actin (Fragment)
OS Gallus gallus (Chicken)
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Archosauria; Aves; Neognathae; Galliformes; Phasianinae;
OC Gallus.
OX NCBI_TaxID=9031;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=84223949; PubMed=6729461;
RA Paterson B.M., Eldridge J.D.;
RT "alpha-Cardiac actin is the major sarcomeric isoform expressed in
RT embryonic avian skeletal muscle.";
RL Science 224:1436-1438(1984).
CC -!- FUNCTION: Actins are highly conserved proteins that are involved
CC in various types of cell motility and are ubiquitously expressed
CC in all eukaryotic cells (By similarity).
CC -!- SIMILARITY: Belongs to the actin family.
DR EMBL; K02173; AAA98513.1; -
DR PIR; C43616; C43616.
DR HSSP; P02577; INMI.
DR GO; GO:0005884; C:actin filament; IEA.
DR GO; GO:0003774; F:motor activity; IEA.
DR GO; GO:0005200; F:structural constituent of cytoskeleton; IEA.
DR Pfam; PF00022; Actin; 1.
DR PRINTS; PR00190; ACTIN.
DR PROSITE; PS00406; ACTINS_1; 1.
KW Structural protein.
FT NON TER 86
SQ SEQUENCE 86 AA; 9351 MW; A55285196A328B6E CRC64;

Query Match 100.0%; Score 82; DB 2; Length 86;
Best Local Similarity 100.0%; Pred. No. 4.5e-06;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```





```
QY 1 GRPRHQGVWVGWQK 15
DB 37 GRPRHQGVWVGWQK 51

RESULT 12
Q902B4 PRELIMINARY; PRT; 99 AA.
AC Q902B4;
DT 01-MAY-2000 (TRENBLrel. 13, Created)
DT 01-MAY-2000 (TRENBLrel. 13, Last sequence update)
DE 01-JUN-2003 (TRENBLrel. 24, Last annotation update)
DE Cytoplasmic actin (Fragment).
OS Cavia porcellus (Guinea pig).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Hystriognathi; Caviidae; Cavia.
OX NCBI_TaxID=10141;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Adrenal gland;
RA Buelow H.E., Bernhardt R.;
RL Submitted (OCT-1999) to the EMBL/GenBank/DBJ databases.
CC -!- FUNCTION: Actins are highly conserved proteins that are involved
CC in various types of cell motility and are ubiquitously expressed
CC in all eukaryotic cells (By similarity).
CC -!- SIMILARITY: Belongs to the actin family.
DR EMBL; AF191277; AAF13923.1; -.
DR HSSP; P02577; INM1.
DR GO; GO:0005884; C:actin filament; IEA.
DR GO; GO:0003774; F:motor activity; IEA.
DR GO; GO:0005200; F:structural constituent of cytoskeleton; IEA.
DR InterPro; IPR004001; Actin.
DR Pfam; PF00022; Actin; 1.
DR PRINTS; PR00190; ACTIN.
DR PROSITE; PS00406; ACTINS_1; 1.
DR Structural protein.
KW NON_TER 99
SQ SEQUENCE 99 AA; 10973 MW; 0562B645E8DAF17B CRC64;

Query Match 100.0%; Score 82; DB 2; Length 99;
Best Local Similarity 100.0%; Pred. No. 5.1e-06;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GRPRHQGVWVGWQK 15
DB 36 GRPRHQGVWVGWQK 50

RESULT 13
Q86SB7 PRELIMINARY; PRT; 104 AA.
AC Q86SB7;
DT 01-JUN-2003 (TRENBLrel. 24, Created)
DT 01-JUN-2003 (TRENBLrel. 24, Last sequence update)
DT 01-MAR-2004 (TRENBLrel. 26, Last annotation update)
DE Actin (Fragment).
GN Name=ACT;
OS Gryllus bimaculatus (Two-spotted cricket).
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC Neoptera; Orthopteroidea; Orthoptera; Ensifera; Gryllidae; Gryllinae;
OC Gryllus.
OX NCBI_TaxID=6999;
RN [1]
RP SEQUENCE FROM N.A.
RA Zhang H., Shimmyo Y., Hirose A., Mito T., Inoue Y., Ohuchi H.,
RA Loukeris T.G., Naji S.;
RT "Extrachromosomal transposition of the transposable element Minos
RT occurs in embryos of the cricket Gryllus bimaculatus.";
RL Dev. Growth Differ. 0:0-0(2002).
CC -!- FUNCTION: Actins are highly conserved proteins that are involved
CC in various types of cell motility and are ubiquitously expressed
CC in all eukaryotic cells (By similarity).

-!- SIMILARITY: Belongs to the actin family.
DR EMBL; AB087882; BAC55093.1; -.
DR HSSP; P02577; INM1.
DR GO; GO:0005884; C:actin filament; IEA.
DR GO; GO:0003774; F:motor activity; IEA.
DR GO; GO:0005200; F:structural constituent of cytoskeleton; IEA.
DR InterPro; IPR004001; Actin.
DR InterPro; IPR004000; Actin_like.
DR Pfam; PF00022; Actin; 1.
DR PRINTS; PR00190; ACTIN.
DR PROSITE; PS00406; ACTINS_1; 1.
DR Structural protein.
FT NON_TER 104
SQ SEQUENCE 104 AA; 11511 MW; 332777A2B08669B0 CRC64;

Query Match 100.0%; Score 82; DB 2; Length 104;
Best Local Similarity 100.0%; Pred. No. 5.3e-06;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GRPRHQGVWVGWQK 15
DB 37 GRPRHQGVWVGWQK 51

RESULT 14
O18548 PRELIMINARY; PRT; 108 AA.
AC O18548;
DT 01-JAN-1998 (TRENBLrel. 05, Created)
DT 01-JAN-1998 (TRENBLrel. 05, Last sequence update)
DT 01-JUN-2003 (TRENBLrel. 24, Last annotation update)
DE Cytoplasmic actin LvC4 (Fragment).
GN Name=LvC4;
OS Lytechinus variegatus (Sea urchin).
OC Eukaryota; Metazoa; Echinodermata; Eleutherozoa; Echinozoa;
OC Echinoidea; Euechinoidea; Echinacea; Temnopleuroidea; Toxopneustidae;
OC Lytechinus.
OX NCBI_TaxID=7654;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=97333945; PubMed=9190067;
RA Kissinger J.C., Hahn J.H., Raff R.A.;
RT "Rapid evolution in a conserved gene family. Evolution of the actin
RT gene family in the sea urchin genus Helicidaris and related genera.";
RL Mol. Biol. Evol. 14:654-665(1997).
CC -!- FUNCTION: Actins are highly conserved proteins that are involved
CC in various types of cell motility and are ubiquitously expressed
CC in all eukaryotic cells (By similarity).
-!- SIMILARITY: Belongs to the actin family.
DR EMBL; U82542; AAB66299.1; -.
DR HSSP; P10983; 1D4X.
DR GO; GO:0005884; C:actin filament; IEA.
DR GO; GO:0003774; F:motor activity; IEA.
DR GO; GO:0005200; F:structural constituent of cytoskeleton; IEA.
DR InterPro; IPR004001; Actin.
DR InterPro; IPR004000; Actin_like.
DR Pfam; PF00022; Actin; 1.
DR PRINTS; PR00190; ACTIN.
DR PROSITE; PS00406; ACTINS_1; 1.
DR PROSITE; PS01132; ACTINS_ACT_LIKE; 1.
DR Structural protein.
FT NON_TER 108
SQ SEQUENCE 108 AA; 12094 MW; F670CB8A6CB3DAE CRC64;

Query Match 100.0%; Score 82; DB 2; Length 108;
Best Local Similarity 100.0%; Pred. No. 5.5e-06;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GRPRHQGVWVGWQK 15
DB 23 GRPRHQGVWVGWQK 37
```

```

RESULT 15
OI8550 PRELIMINARY; PRT; 108 AA.
ID OI8550
AC OI8550;
DT 01-JAN-1998 (TrEMBLrel. 05, Created)
DT 01-JAN-1998 (TrEMBLrel. 05, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE Cytoplasmic actin LVC2 (fragment).
GN Name=LVC2;
OS Lytechinus variegatus (Sea urchin).
OC Eukaryota; Metazoa; Echinodermata; Eleutherozoa; Echinozoa;
OC Echinoidea; Euechinoidea; Echinacea; Temnopneustidae;
OC Lytechinus;
OX NCBI_TaxID=7654;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=97333945; PubMed=9190067;
RT Kissinger J.C., Hahn J.H., Raff R.A.;
RT "Rapid evolution in a conserved gene family. Evolution of the actin
RT gene family in the sea urchin genus Helicidaris and related genera.";
RL Mol. Biol. Evol. 14:654-665(1997).
CC -!- FUNCTION: Actins are highly conserved proteins that are involved
CC in various types of cell motility and are ubiquitously expressed
CC in all eukaryotic cells (By similarity).
CC -!- SIMILARITY: belongs to the actin family.
CC EMBL; U02544; AAB6301.1; -.
CC HSP; P10983; IDAX.
DR GO; GO:0005884; C:actin filament; IEA.
DR GO; GO:0003774; F:motor activity; IEA.
DR GO; GO:0005200; F:structural constituent of cytoskeleton; IEA.
DR InterPro; IPR004001; Actin.
DR InterPro; IPR004000; Actin_like.
DR Pfam; PF00022; Actin; 1.
DR PRINTS; PR00190; ACTIN.
DR PROSITE; PS00406; ACTINS 1; 1.
DR PROSITE; PS01132; ACTINS_ACT_LIKE; 1.
KW Structural protein.
FT NON_TER 1
FT NON_TER 108
SQ SEQUENCE 108 AA; 12097 MW; F670D62101164AEA CRC64;

Query Match 100.0%; Score 82; DB 2; Length 108;
Best Local Similarity 100.0%; Pred. No. 5.5e-06;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GRPRHQGVNMGQK 15
DB 23 GRPRHQGVNMGQK 37

Search completed: April 8, 2005, 12:03:06
Job time : 57.8667 secs

```

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: April 8, 2005, 09:58:56 ; Search time 62.4 Seconds  
(without alignments)  
92.971 Million cell updates/sec

Title: US-09-423-351C-4

Perfect score: 81

Sequence: 1 GGTTMPGIADRMQK 15

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 2105692 seqs, 386760381 residues

Total number of hits satisfying chosen parameters: 2105692

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : A\_Geneseq\_16Dec04:\*

- 1: Geneseqp1980s:\*
- 2: Geneseqp1990s:\*
- 3: Geneseqp2000s:\*
- 4: Geneseqp2001s:\*
- 5: Geneseqp2002s:\*
- 6: Geneseqp2003as:\*
- 7: Geneseqp2003bs:\*
- 8: Geneseqp2004s:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	81	100.0	15	2	AAW92530 Beta-acti
2	81	100.0	86	4	ABG08617 Novel hum
3	81	100.0	100	4	AAW81404 Human hae
4	81	100.0	100	4	AAW81907 Human hae
5	81	100.0	110	5	ABP42328 Human ova
6	81	100.0	118	4	AAU30705 Novel hum
7	81	100.0	145	5	ABP43127 Human ova
8	81	100.0	147	4	AAW80875 Human hae
9	81	100.0	147	4	AAW81886 Human hae
10	81	100.0	147	4	AAW81332 Human hae
11	81	100.0	148	3	AAW59047 Breast an
12	81	100.0	149	8	ABO84771 Human can
13	81	100.0	196	6	ABU70816 Human adi
14	81	100.0	214	8	ADR09052 Human pro
15	81	100.0	258	4	ABG00383 Novel hum
16	81	100.0	274	6	ABU70549 Human adi
17	81	100.0	294	7	AD163068 Human apo
18	81	100.0	297	4	ABG15100 Novel hum
19	81	100.0	332	7	AD163011 Human apo
20	81	100.0	334	6	ABP98860 Human str
21	81	100.0	342	7	ADM05323 Human pro
22	81	100.0	374	3	AAU78101 Misgurnus
23	81	100.0	374	3	AAU12985 Human bet
24	81	100.0	375	1	AAW61532 Sequence
25	81	100.0	375	2	AAW50328 Drug resi

26	81	100.0	375	2	AAW59176
27	81	100.0	375	3	AAW15017
28	81	100.0	375	3	AAW94569
29	81	100.0	375	5	ABW77395
30	81	100.0	375	6	ABW64271
31	81	100.0	375	6	ABW62328
32	81	100.0	375	6	ABW04830
33	81	100.0	375	7	ADW85212
34	81	100.0	375	7	ADW61174
35	81	100.0	375	7	ADW30525
36	81	100.0	375	7	ADW163062
37	81	100.0	375	7	ADW162970
38	81	100.0	375	7	ADW163040
39	81	100.0	375	8	ADW13002
40	81	100.0	375	8	ADW78489
41	81	100.0	375	8	ADW14103
42	81	100.0	375	8	ADW12391
43	81	100.0	375	8	ADW26098
44	81	100.0	375	8	ABW84772
45	81	100.0	375	8	ABW80841

ALIGNMENTS

RESULT 1  
AAW92530  
ID AAW92530 standard; peptide; 15 AA.  
XX  
AC AAW92530;  
XX  
DT 26-APR-1999 (first entry)  
XX  
DE Beta-actin reference peptide substrate #4.  
XX

Peptide substrate; CCT; eukaryotic type II chaperonin complex; cyclin; binding agent; substrate-binding site; SBS; substrate folding; actin; tubulin; treatment; cancer; anticancer drug; viral infection; screening; reduced toxicity.

XX Synthetic.

XX WO9853322-A1.

XX 26-NOV-1998.

XX 22-MAY-1998; 98WO-GB001485.

XX 23-MAY-1997; 97GB-00010762.

XX (CANC-) INST CANCER RES ROYAL CANCER HOSPITAL.

XX Willison K, Hynes G, Liou AK;

XX WPI; 1999-070162/06.

Identifying specific binding agents for substrate binding site in CCT chaperonin complex - also new peptide binding agents and their mimetics, and peptides containing a specific CCT binding site, used for treating cancer.

Disclosure; Fig 10; 97pp; English.

This invention describes a method which uses the CCT (eukaryotic type II chaperonin) complex or part of it, for identifying a binding agent that can occupy a substrate-binding site (SBS) on the CCT complex. By binding to the CCT complex, the binding agents block an SBS so that biological activity of the CCT complex is affected, particularly its ability to fold substrates such as actin, tubulin and cyclin. The binding agents are useful for treatment of cancer, particularly when used in combination with an anticancer drug, or viral infections. Nucleic acid fragments are used to screen for agents, e.g. binding agents that modulate interaction between the CCT complex and a protein that is to be folded. The binding

CC agents may target cells that are actively synthesising tubulin etc.  
 CC (unlike known microtubule-stabilising agents that affect all cells), so  
 CC should have reduced toxicity for normal cells. AAW92527-W92541 are  
 CC peptide substrates used in the method of the invention  
 XX

XX Sequence 15 AA;  
 Query Match 100.0%; Score 81; DB 2; Length 15;  
 Best Local Similarity 100.0%; Pred. No. 3.4e-07;  
 Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 Qy 1 GGTMYPGIADRMQK 15  
 |||||  
 Db 1 GGTMYPGIADRMQK 15  
 |||||

RESULT 2  
 ABOG08617  
 ID ABOG08617 standard; protein; 86 AA.  
 XX  
 AC ABOG08617;  
 XX  
 DT 13-FEB-2002 (first entry)  
 DE  
 DE Novel human diagnostic protein #8608.

XX Human; chromosome mapping; gene mapping; gene therapy; forensic;  
 KW food supplement; medical imaging; diagnostic; genetic disorder.  
 XX  
 OS Homo sapiens.  
 XX  
 PN WO200175067-A2.  
 XX  
 PD 11-OCT-2001.

XX 30-MAR-2001; 2001WO-US008631.  
 XX 31-MAR-2000; 2000US-00540217.  
 PR 23-AUG-2000; 2000US-00649167.  
 XX (HYSE-) HYSEQ INC.  
 XX Drmanac RT, Liu C, Tang YT;  
 XX WPI; 2001-639362/73.  
 DR N-PSDB; AAS72804.

XX New isolated polynucleotide and encoded polypeptides, useful in  
 PT diagnostics, forensics, gene mapping, identification of mutations  
 PT responsible for genetic disorders or other traits and to assess  
 PT biodiversity.

XX Claim 20; SEQ ID NO 38976; 103pp; English.

XX The invention relates to isolated polynucleotide (I) and polypeptide (II)  
 CC sequences. (I) is useful as hybridisation probes, polymerase chain  
 CC reaction (PCR) primers, oligomers, and for chromosome and gene mapping,  
 CC and in recombinant production of (II). The polynucleotides are also used  
 CC in diagnostics as expressed sequence tags for identifying expressed  
 CC genes. (I) is useful in gene therapy techniques to restore normal  
 CC activity of (II) or to treat disease states involving (II). (II) is  
 CC useful for generating antibodies against it, detecting or quantitating a  
 CC polypeptide in tissue, as molecular weight markers and as a food  
 CC supplement. (II) and its binding partners are useful in medical imaging  
 CC of sites expressing (II). (I) and (II) are useful for treating disorders  
 CC involving aberrant protein expression or biological activity. The  
 CC polypeptide and polynucleotide sequences have applications in  
 CC diagnostics, forensics, gene mapping, identification of mutations  
 CC responsible for genetic disorders or other traits to assess biodiversity  
 CC and to produce other types of data and products dependent on DNA and  
 CC amino acid sequences. ABOG0010-ABG30377 represent novel human diagnostic  
 CC amino acid sequences of the invention. Note: The sequence data for this  
 CC patent did not appear in the printed specification, but was obtained in

CC electronic format directly from WIPO at  
 CC ftp.wipo.int/pub/published\_pct\_sequences  
 XX  
 SQ Sequence 86 AA;

Query Match 100.0%; Score 81; DB 4; Length 86;  
 Best Local Similarity 100.0%; Pred. No. 2.2e-06;  
 Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 Qy 1 GGTMYPGIADRMQK 15  
 |||||  
 Db 12 GGTMYPGIADRMQK 26  
 |||||

RESULT 3  
 AAM81404  
 ID AAM81404 standard; protein; 100 AA.  
 XX  
 AC AAM81404;  
 XX  
 DT 13-NOV-2001 (first entry)  
 DE  
 DE Human haematological malignancy-related antigen #1102.

XX Human; cytostatic; vascular; gene therapy; vaccine; lymphoma;  
 KW haematological malignancy; antigen; chronic lymphocytic leukaemia;  
 KW follicular lymphoma; Hodgkin's lymphoma; non-Hodgkin's lymphoma.  
 XX  
 OS Homo sapiens.  
 XX  
 PN WO200164886-A2.  
 XX  
 PD 07-SEP-2001.

XX 01-MAR-2001; 2001WO-US007272.  
 XX 01-MAR-2000; 2000US-0186126P.  
 PR 17-MAR-2000; 2000US-0190479P.  
 PR 27-APR-2000; 2000US-0200545P.  
 PR 28-APR-2000; 2000US-0200303P.  
 PR 28-APR-2000; 2000US-0200779P.  
 PR 01-MAY-2000; 2000US-0200999P.  
 PR 04-MAY-2000; 2000US-0202084P.  
 PR 22-MAY-2000; 2000US-0206201P.  
 PR 14-JUL-2000; 2000US-0218950P.  
 PR 03-AUG-2000; 2000US-0222903P.  
 PR 04-AUG-2000; 2000US-0223416P.  
 PR 07-AUG-2000; 2000US-0223378P.  
 XX (CORI-) CORIXA CORP.

XX Gaiger A, Algate PA, Mannion J;  
 XX WPI; 2001-514842/56.

XX Compositions and methods for the detection of hematological malignancies,  
 PT e.g. chronic lymphocytic leukemia, lymphoma, follicular lymphoma and  
 PT Hodgkin's and T/B cell non-Hodgkin's lymphoma.

XX Claim 1; Page 953; 1252pp; English.  
 XX The present invention relates to compositions and methods for the  
 CC detection, diagnosis and therapy of haematological malignancies. The  
 CC present sequence is the protein sequence of a human haematological  
 CC malignancy related antigen. The methods of the present invention comprise  
 CC detecting the presence of haematological malignancy related antigen(s) in  
 CC a sample obtained from the patient (an increased level of the  
 CC polypeptide, compared to an unaffected individual, is indicative of an  
 CC increased risk). Haematological malignancies which can be treated using  
 CC the present invention are chronic lymphocytic leukaemia, lymphoma,  
 CC follicular lymphoma, Hodgkin's lymphoma, non-Hodgkin's lymphoma and T/B  
 CC cell non-Hodgkin's lymphoma

SQ Sequence 100 AA;  
 Query Match 100.0%; Score 81; DB 4; Length 100;  
 Best Local Similarity 100.0%; Pred. No. 2.6e-06;  
 Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGTMYPGIADRMQK 15  
 DB 46 GGTMYPGIADRMQK 60

RESULT 4  
 AAM81907  
 ID AAM81907 standard; protein; 100 AA.  
 XX AC AAM81907;  
 XX DT 13-NOV-2001 (first entry)  
 XX DE Human haematological malignancy-related antigen #1605.  
 XX KW Human; cytostatic; vascular; gene therapy; vaccine; lymphoma;  
 KW haematological malignancy; antigen; chronic lymphocytic leukaemia;  
 KW follicular lymphoma; Hodgkin's lymphoma; non-Hodgkin's lymphoma.  
 XX OS Homo sapiens.  
 XX PN WO200164886-A2.  
 XX PD 07-SEP-2001.  
 XX PF 01-MAR-2001; 2001WO-US007272.  
 XX PR 01-MAR-2000; 2000US-0186126P.  
 PR 17-MAR-2000; 2000US-0190479P.  
 PR 27-APR-2000; 2000US-0200545P.  
 PR 28-APR-2000; 2000US-0200303P.  
 PR 28-APR-2000; 2000US-0200779P.  
 PR 01-MAY-2000; 2000US-0200999P.  
 PR 04-MAY-2000; 2000US-0202084P.  
 PR 22-MAY-2000; 2000US-0208201P.  
 PR 14-JUL-2000; 2000US-0218950P.  
 PR 03-AUG-2000; 2000US-0222903P.  
 PR 04-AUG-2000; 2000US-0223416P.  
 PR 07-AUG-2000; 2000US-0223378P.  
 XX PA (CORI-) CORIXA CORP.  
 XX PI Gaiger A, Algate PA, Mannion J;  
 XX DR WPI; 2001-514842/56.  
 XX PS Claim 1; Page 1149-1150; 1252pp; English.  
 XX CC The present invention relates to compositions and methods for the  
 CC detection, diagnosis and therapy of haematological malignancies. The  
 CC present sequence is the protein sequence of a human haematological  
 CC malignancy related antigen. The methods of the present invention comprise  
 CC detecting the presence of haematological malignancy related antigen(s) in  
 CC a sample obtained from the patient (an increased level of the  
 CC polypeptide, compared to an unaffected individual, is indicative of an  
 CC increased risk). Haematological malignancies which can be treated using  
 CC the present invention are chronic lymphocytic leukaemia, lymphoma,  
 CC follicular lymphoma, Hodgkin's lymphoma, non-Hodgkin's lymphoma and T/B  
 CC cell non-Hodgkin's lymphoma  
 XX SQ Sequence 100 AA;  
 Query Match 100.0%; Score 81; DB 4; Length 100;

Best Local Similarity 100.0%; Pred. No. 2.6e-06;  
 Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGTMYPGIADRMQK 15  
 DB 46 GGTMYPGIADRMQK 60

RESULT 5  
 ABP42928  
 ID ABP42928 standard; protein; 110 AA.  
 XX AC ABP42928;  
 XX DT 22-AUG-2002 (first entry)  
 XX DE Human ovarian antigen HPRS46, SEQ ID NO:4060.  
 XX KW Human; ovarian antigen; ovary; ovarian; breast; cancer; tumour;  
 KW ovarian cancer; breast cancer; tumour; reproductive system disorder;  
 KW infertility; pregnancy disorder; anovulation; polycystic ovary syndrome;  
 KW PCOS; ovarian cyst; dysmenorrhoea; endocrine disorder; infection;  
 KW inflammatory condition; immune disorder; blood disorder;  
 KW cardiovascular disorder; respiratory disorder; neurological disorder;  
 KW gastrointestinal disorder; urinary system disorder; drug screening;  
 KW gene therapy; chromosome mapping; forensic analysis;  
 KW antibody preparation; cytostatic; immunomodulatory; neuroprotective;  
 KW antiinflammatory; gynaecological; reproductive.  
 XX OS Homo sapiens.  
 XX PN WO200200677-A1.  
 XX PD 03-JAN-2002.  
 XX PF 07-JUN-2001; 2001WO-US018569.  
 XX PR 07-JUN-2000; 2000US-0209467P.  
 XX PA (HUMA-) HUMAN GENOME SCI INC.  
 XX PI Birse CE, Rosen CA;  
 XX DR WPI; 2002-147878/19.  
 XX DR N-PSDB; ABQ56005.  
 XX PT Isolated nucleic acid molecules encoding novel ovarian polypeptides,  
 PT useful in the prevention, treatment and diagnosis of cancer (e.g. ovarian  
 PT cancer), immune disorders, cardiovascular disorders and neurological  
 PT diseases.  
 XX PS Claim 11; SEQ ID NO 4060; 2922pp; English.  
 XX CC The invention relates to 2175 novel human ovarian antigens (ABP41054-  
 CC ABP43228) and to cDNAs encoding them (ABQ54131-ABQ56305), and also  
 CC encompasses polypeptides 90% identical and polynucleotides 95% identical  
 CC to the sequences of the invention. The invention additionally relates to  
 CC recombinant vectors and host cells comprising human ovarian antigen  
 CC polynucleotides, antibodies against human ovarian antigens, and the use  
 CC of ovarian antigen polynucleotides and polypeptides in diagnosing,  
 CC treating, prognosing or preventing various ovary and/or breast-related  
 CC disorders. Such conditions include ovarian cancer and breast cancer, and  
 CC metastatic tumours of ovarian or breast origin, reproductive system  
 CC disorders (e.g., infertility, disorders of pregnancy, anovulation,  
 CC polycystic ovary syndrome, ovarian cysts, and dysmenorrhoea), endocrine  
 CC disorders, infections (e.g., chlamydia, HIV, toxoplasmosis, and toxic  
 CC shock syndrome), inflammatory conditions (e.g., mastitis, oophoritis and  
 CC vaginitis), immune disorders (e.g., congenital and acquired  
 CC immunodeficiencies, autoimmune oophoritis, systemic lupus erythematosus),  
 CC blood-related disorders (e.g., anaemia), cardiovascular disorders,  
 CC respiratory disorders, neurological disorders, gastrointestinal disorders  
 CC and urinary system disorders. Ovarian antigen polypeptides and  
 CC polynucleotides may also be used in screening for compounds which

CC modulate ovarian antigen expression or activity. The polynucleotides may  
CC further be used for gene therapy, chromosome mapping, in the  
CC identification of individuals and in forensic analysis, and the  
CC polypeptides may be used as food additives or to prepare antibodies  
CC useful in disease diagnosis, drug targeting and phenotyping. The present  
CC sequence represents a human ovarian antigen of the invention. Note: The  
CC sequence data for this patent did not form part of the printed  
CC specification, but was obtained in electronic format directly from WIPO  
CC at ftp.wipo.int/pub/published\_pct\_sequences  
XX  
XX  
SQ Sequence 110 AA;

Query Match 100.0%; Score 81; DB 5; Length 110;  
Best Local Similarity 100.0%; Pred. No. 2.9e-06;  
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGTMYPGIADRMQK 15  
|||||  
Db 36 GGTMYPGIADRMQK 50

RESULT 6  
AAU30705  
ID AAU30705 standard; protein; 118 AA.

XX  
XX  
AC AAU30705;

DT 18-DEC-2001 (first entry)

XX  
XX  
DE Novel human secreted protein #1196.

XX Human; vaccination; gene therapy; nutritional supplement;  
KW stem cell proliferation; hematopoiesis; nerve tissue regeneration;  
KW immune suppression; immune stimulation; anti-inflammatory; leukaemia.

XX Homo sapiens.

XX WO200179449-A2.

XX 25-OCT-2001.

XX 16-APR-2001; 2001WO-US008656.

XX 18-APR-2000; 2000US-00552929.

XX 26-JAN-2001; 2001US-00770160.

XX (HYSE-) HYSEQ INC.

XX Tang YT, Liu C, Drmanac RT;

XX WPI; 2001-611725/70.

XX Nucleic acids encoding a range of human polypeptides, useful in genetic  
PT vaccination, testing and therapy.

XX Claim 20; Page 333; 765pp; English.

XX The invention relates to novel human secreted polypeptides. The  
CC polypeptides and antibodies to the polypeptides are useful for  
CC determining the presence of or predisposition to a disease associated  
CC with altered levels of polypeptide. The polypeptides are also useful for  
CC identifying agents (agonists and antagonists) that bind to them. Cells  
CC expressing the proteins are useful for identifying a therapeutic agent  
CC for use in treatment of a pathology related to aberrant expression or  
CC physiological interactions of the polypeptide. Vectors comprising the  
CC nucleic acids encoding the polypeptides and cells genetically engineered  
CC to express them are also useful for producing the proteins. The proteins  
CC are useful in genetic vaccination, testing and therapy, and can be used  
CC as nutritional supplements. They may be used to increase stem cell  
CC proliferation, to regulate haematopoiesis; and in bone, cartilage, tendon  
CC and/or nerve tissue growth or regeneration; immune suppression and/or  
CC stimulation; as anti-inflammatory agents; and in treatment of leukaemias.  
CC AAU29510-AAU33304 represent the amino acid sequences of novel human

CC secreted proteins of the invention  
XX  
SQ Sequence 118 AA;

Query Match 100.0%; Score 81; DB 4; Length 118;  
Best Local Similarity 100.0%; Pred. No. 3.1e-06;  
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGTMYPGIADRMQK 15  
|||||  
Db 42 GGTMYPGIADRMQK 56

RESULT 7

ABP43127  
ID ABP43127 standard; protein; 145 AA.

XX  
XX  
AC ABP43127;

XX 22-AUG-2002 (first entry)

XX Human ovarian antigen HVBD91, SEQ ID NO:4259.

XX Human; ovarian antigen; ovary; ovarian; breast; cancer; tumour;  
KW ovarian cancer; breast cancer; tumour; reproductive system disorder;  
KW infertility; pregnancy disorder; anovulation; polycystic ovary syndrome;  
KW PCOS; ovarian cyst; dysmenorrhoea; endocrine disorder; infection;  
KW inflammatory condition; immune disorder; blood disorder;  
KW cardiovascular disorder; respiratory disorder; neurological disorder;  
KW gastrointestinal disorder; urinary system disorder; drug screening;  
KW gene therapy; chromosome mapping; forensic analysis;  
KW antibody preparation; cytostatic; immunomodulatory; neuroprotective;  
KW antiinflammatory; gynaecological; reproductive.

XX Homo sapiens.

XX WO200200677-A1.

XX 03-JAN-2002.

XX 07-JUN-2001; 2001WO-US018569.

XX 07-JUN-2000; 2000US-0209467P.

XX (HUMA-) HUMAN GENOME SCI INC.

XX Birse CE, Rosen CA;

XX WPI; 2002-147878/19.

XX N-PSDB; ABQ56204.

XX Isolated nucleic acid molecules encoding novel ovarian polypeptides,  
PT useful in the prevention, treatment and diagnosis of cancer (e.g. ovarian  
PT cancer), immune disorders, cardiovascular disorders and neurological  
PT diseases.

XX Claim 11; SEQ ID NO 4259; 2922pp; English.

XX The invention relates to 2175 novel human ovarian antigens (ABP41054-  
CC ABP43228) and to cDNAs encoding them (ABQ54131-ABQ56305), and also  
CC encompasses polypeptides 90% identical and polynucleotides 95% identical  
CC to the sequences of the invention. The invention additionally relates to  
CC recombinant vectors and host cells comprising human ovarian antigen  
CC polynucleotides, antibodies against human ovarian antigens, and the use  
CC of ovarian antigen polynucleotides and polypeptides in diagnosing,  
CC treating, prognosing or preventing various ovary and/or breast-related  
CC disorders. Such conditions include ovarian cancer and breast cancer, and  
CC metastatic tumours of ovarian or breast origin, reproductive system  
CC disorders (e.g., infertility, disorders of pregnancy, anovulation,  
CC polycystic ovary syndrome, ovarian cysts, and dysmenorrhoea), endocrine  
CC disorders, infections (e.g., chlamydia, HIV, toxoplasmosis), and toxic  
CC shock syndrome), inflammatory conditions (e.g., mastitis, oophoritis and  
CC vaginitis), immune disorders (e.g., congenital and acquired

CC immunodeficiencies, autoimmune oophoritis, systemic lupus erythematosus),  
 CC blood-related disorders (e.g., anemia), cardiovascular disorders,  
 CC respiratory disorders, neurological disorders, gastrointestinal disorders  
 CC and urinary system disorders. Ovarian antigen polypeptides and  
 CC polynucleotides may also be used in screening for compounds which  
 CC modulate ovarian antigen expression or activity. The polynucleotides may  
 CC further be used for gene therapy, chromosome mapping, in the  
 CC identification of individuals and in forensic analysis, and the  
 CC polypeptides may be used as food additives or to prepare antibodies  
 CC useful in disease diagnosis, drug targeting and phenotyping. The present  
 CC sequence represents a human ovarian antigen of the invention. Note: The  
 CC sequence data for this patent did not form part of the printed  
 CC specification, but was obtained in electronic format directly from WIPO  
 CC at ftp.wipo.int/pub/published\_pct\_sequences  
 XX  
 SQ Sequence 145 AA;

Query Match 100.0%; Score 81; DB 5; Length 145;  
 Best Local Similarity 100.0%; Pred. No. 3.9e-06;  
 Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGTMYPGIADRMQK 15  
 |||||  
 DB 71 GGTMYPGIADRMQK 85

RESULT 8  
 AAM80875  
 ID AAM80875 standard; protein; 147 AA.  
 XX AC AAM80875;  
 XX DT 13-NOV-2001 (first entry)  
 XX DE Human haematological malignancy-related antigen #573.  
 XX KW Human; cytostatic; vascular; gene therapy; vaccine; lymphoma;  
 KW haematological malignancy; antigen; chronic lymphocytic leukaemia;  
 KW follicular lymphoma; Hodgkin's lymphoma; non-Hodgkin's lymphoma.  
 XX OS Homo sapiens.

XX WO200164886-A2.  
 XX PD 07-SEP-2001.  
 XX PF 01-MAR-2001; 2001WO-US007272.  
 XX PR 01-MAR-2000; 2000US-0186126P.  
 XX PR 17-MAR-2000; 2000US-0190479P.  
 XX PR 27-APR-2000; 2000US-0200545P.  
 XX PR 28-APR-2000; 2000US-0200303P.  
 XX PR 01-MAY-2000; 2000US-0200779P.  
 XX PR 04-MAY-2000; 2000US-0202084P.  
 XX PR 22-MAY-2000; 2000US-0206201P.  
 XX PR 03-AUG-2000; 2000US-0222903P.  
 XX PR 07-AUG-2000; 2000US-0223378P.

(CORI-) CORIXA CORP.  
 XX Gaiger A, Algate PA, Mannion J;  
 XX WPI; 2001-514842/56.  
 XX Compositions and methods for the detection of hematological malignancies,  
 PT e.g. chronic lymphocytic leukemia, lymphoma, follicular lymphoma and  
 PT Hodgkin's and T/B cell non-Hodgkin's lymphoma.  
 XX Claim 1; Page 743-744; 1252pp; English.

CC The present invention relates to compositions and methods for the  
 CC detection, diagnosis and therapy of haematological malignancies. The  
 CC present sequence is the protein sequence of a human haematological  
 CC malignancy related antigen. The methods of the present invention comprise  
 CC detecting the presence of haematological malignancy related antigen(s) in  
 CC a sample obtained from the patient (an increased level of the  
 CC polypeptide, compared to an unaffected individual, is indicative of an  
 CC increased risk). Haematological malignancies which can be treated using  
 CC the present invention are chronic lymphocytic leukaemia, lymphoma,  
 CC follicular lymphoma, Hodgkin's lymphoma, non-Hodgkin's lymphoma and T/B  
 CC cell non-Hodgkin's lymphoma  
 XX  
 SQ Sequence 147 AA;

Query Match 100.0%; Score 81; DB 4; Length 147;  
 Best Local Similarity 100.0%; Pred. No. 4e-06;  
 Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGTMYPGIADRMQK 15  
 |||||  
 DB 73 GGTMYPGIADRMQK 87

RESULT 9  
 AAM81886  
 ID AAM81886 standard; protein; 147 AA.  
 XX AC AAM81886;  
 XX DT 13-NOV-2001 (first entry)  
 XX DE Human haematological malignancy-related antigen #1584.

XX KW Human; cytostatic; vascular; gene therapy; vaccine; lymphoma;  
 KW haematological malignancy; antigen; chronic lymphocytic leukaemia;  
 KW follicular lymphoma; Hodgkin's lymphoma; non-Hodgkin's lymphoma.  
 XX OS Homo sapiens.

XX WO200164886-A2.  
 XX PD 07-SEP-2001.  
 XX PF 01-MAR-2001; 2001WO-US007272.  
 XX PR 01-MAR-2000; 2000US-0186126P.  
 XX PR 17-MAR-2000; 2000US-0190479P.  
 XX PR 27-APR-2000; 2000US-0200545P.  
 XX PR 28-APR-2000; 2000US-0200303P.  
 XX PR 01-MAY-2000; 2000US-0200779P.  
 XX PR 04-MAY-2000; 2000US-0202084P.  
 XX PR 22-MAY-2000; 2000US-0206201P.  
 XX PR 14-JUL-2000; 2000US-0218950P.  
 XX PR 03-AUG-2000; 2000US-0222903P.  
 XX PR 04-AUG-2000; 2000US-0223416P.  
 XX PR 07-AUG-2000; 2000US-0223378P.

(CORI-) CORIXA CORP.  
 XX Gaiger A, Algate PA, Mannion J;  
 XX WPI; 2001-514842/56.  
 XX Compositions and methods for the detection of hematological malignancies,  
 PT e.g. chronic lymphocytic leukemia, lymphoma, follicular lymphoma and  
 PT Hodgkin's and T/B cell non-Hodgkin's lymphoma.

XX Claim 1; Page 1141; 1252pp; English.  
 XX The present invention relates to compositions and methods for the  
 CC detection, diagnosis and therapy of haematological malignancies. The  
 CC present sequence is the protein sequence of a human haematological





CC of immune disorders e.g. Addison's disease, allergies, autoimmune  
CC haemolytic anaemia, autoimmune thyroiditis, diabetes mellitus, Crohn's  
CC disease, multiple sclerosis, rheumatoid arthritis and ulcerative colitis;  
CC cardiovascular disorders such as myocardial ischaemias; wound healing;  
CC neurological diseases such as cerebral anoxia and epilepsy; and  
CC infectious diseases  
XX  
SQ Sequence 148 AA;

Query Match 100.0%; Score 81; DB 3; Length 148;  
Best Local Similarity 100.0%; Pred. No. 4e-06;  
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGTMYPGIADRMOK 15  
DB 19 GGTMYPGIADRMOK 33

RESULT 12  
ABO84771  
ID ABO84771 standard; protein; 149 AA.

XX AC ABO84771;

XX DT 18-NOV-2004 (first entry)

XX DE Human cancer-associated protein (CAP) HP07-010.

XX KW Human; cancer-associated protein; CAP; cancer; cytostatic.

XX OS Homo sapiens.

XX PN WO2004058146-A2.

XX PD 15-JUL-2004.

XX PF 15-DEC-2003; 2003WO-US040081.

XX PR 17-DEC-2002; 2002US-00322281.

XX PA (SAGR-) SAGRES DISCOVERY INC.

XX PI Morris DW, Malandro MS;

XX DR WPI; 2004-499109/47.

XX DR N-PSDB; ABD33111.

XX PT Novel human cancer associated protein encoded within open reading frame  
XX of cancer associated gene, useful as targets for diagnosing cancer.

XX PS Claim 18; SEQ ID NO 66; 182pp; English.

XX CC The invention relates to cancer-associated proteins (CAP) and the cancer-  
XX associated (CA) nucleic acids encoding them. The invention also relates  
XX to a method for treating cancers involving administering to a patient an  
XX inhibitor of CAP, and a method of screening for anticancer activity in a  
XX potential drug involving providing a cell that expresses a CA gene,  
XX contacting a tissue sample derived from a cancer cell with an anticancer  
XX drug candidate and monitoring the effect of the anticancer drug candidate  
XX on expression of the CA gene. The CAP proteins are useful for detecting  
XX cancer associated with expression of a CAP protein in a test cell sample  
XX and for screening for a bioactive agent capable of modulating the  
XX activity of a CAP protein. The CA nucleic acids are useful for diagnosing  
XX cancer, involving determining the expression of a CA nucleic acid in a  
XX tissue. This sequence represents a human CAP of the invention. Note: The  
XX sequence data for this patent did not form part of the printed  
XX specification, but was obtained in electronic format directly from WIPO  
XX at ftp.wipo.int/pub/published\_pct\_sequences

XX SQ Sequence 149 AA;

Query Match 100.0%; Score 81; DB 8; Length 149;  
Best Local Similarity 100.0%; Pred. No. 4e-06;

Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 GGTMYPGIADRMOK 15  
DB 75 GGTMYPGIADRMOK 89

RESULT 13

ABU70816

ID ABU70816 standard; protein; 196 AA.

XX AC ABU70816;

XX DT 10-JUN-2003 (first entry)

XX DE Human adipocyte Selected Interacting domain, SID, #447.

XX KW Human; prey; adipocyte; SID; selected interacting domain; anorectic;  
XX antidabetic; protein-protein interaction; diabetes;  
XX yeast 2-hybrid assay; metabolic disorder; obesity.

XX OS Homo sapiens.

XX PN WO200286122-A2.

XX PD 31-OCT-2002.

XX PF 14-MAR-2002; 2002WO-EP003768.

XX PR 14-MAR-2001; 2001US-0275734P.

XX PA (HYBR-) HYBRIGENICS.

XX PI Legrain P, Daviet L;

XX DR WPI; 2003-103412/09.

XX DR N-PSDB; ACAS7360.

XX PT New complex between two interacting proteins in adipocyte cells, useful  
XX for identifying selected interacting domains that modulate protein  
XX interactions, or for preventing or treating metabolic disorders such as  
XX obesity or diabetes.

XX PS Claim 6; Page 259-260; 382pp; English.

XX CC The invention relates to a complex between two interacting proteins in  
XX adipocyte cells, given in the specification. The proteins are identified  
XX by selecting a bait protein from a known adipocyte marker and then  
XX performing a yeast 2-hybrid selection to isolate prey proteins encoded by  
XX members of an adipocyte cDNA library. The proteins are designated SID  
XX (RTM) (selected interacting domains) proteins. Also included are a  
XX polynucleotide encoding a polypeptide in the adipocyte cells, a  
XX recombinant host cell expressing at least one of the interacting  
XX polypeptides of the complex, selecting a modulating compound in adipocyte  
XX cells, a SID (RTM) polypeptide comprising any of the 738 amino acid  
XX sequences given in the specification (including its fragment or variant),  
XX a SID (RTM) polynucleotide comprising any of the 738 nucleotide sequences  
XX given in the specification (including its fragment or variant), a vector  
XX comprising the SID (RTM) polynucleotide, a recombinant host cell  
XX comprising the vector, a protein chip comprising the polypeptides and a  
XX record comprising all or part of the data, listed in the specification.  
XX The complex, polypeptides, polynucleotides and compounds are useful for  
XX preventing or treating metabolic disorders such as obesity or diabetes.  
XX The polynucleotides are useful as probes or primers. The complex is  
XX particularly useful for identifying selected interacting domains (SID  
XX (RTM)) for screening drugs that modulate the protein interaction, thus  
XX exhibiting the therapeutic effect. The present sequence represents a SID  
XX (prey) protein of the invention

XX SQ Sequence 196 AA;

Query Match 100.0%; Score 81; DB 6; Length 196;  
Best Local Similarity 100.0%; Pred. No. 5.4e-06;

Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GGTMYPGIADRMQK 15  
 |||||  
 Db 175 GGTMYPGIADRMQK 189

RESULT 14  
 ADR09052  
 ID ADR09052 standard; protein; 214 AA.

XX AC ADR09052;  
 XX DT 04-NOV-2004 (first entry)  
 XX DE Human protein useful for treating neurological disease Seq 2558.

XX KW human; oligo-capping method; diagnostic marker; gene therapy;  
 KW osteoporosis; neurological disease; Alzheimer's disease;  
 KW Parkinson's disease; dementia; short memory; cancer;  
 KW sense or motor function; emotional reaction; fear response; panic;  
 KW osteopathic; neuroprotective; nootropic; antiparkinsonian; cytostatic;  
 KW tranquiliser.

XX OS Homo sapiens.  
 XX FN EP1447413-A2.  
 XX PD 18-AUG-2004.

XX PF 12-FEB-2004; 2004EP-00003145.  
 XX PR 14-FEB-2003; 2003JP-00102207.  
 XX PR 09-MAY-2003; 2003JP-00131452.

XX PA (REAS-) RES ASSOC BIOTECHNOLOGY.

XX PI Isoqai T, Yamamoto J, Nishikawa T, Isono Y, Sugiyama T, Otsuki T;  
 PI Wakamatsu A, Ishii S, Nagai K, Irie R;  
 XX WPI; 2004-583265/57.  
 DR N-PSDB; ADR07096.

XX PT New 1995 cDNA, useful for treating osteoporosis, neurological diseases,  
 PT Alzheimer's diseases, Parkinson's diseases, dementia and various cancers.  
 XX PS Claim 1; SEQ ID NO 2558; 2686pp; English.

XX This invention relates to novel, isolated full length human cDNA  
 CC molecules and the encoded proteins thereof. Specifically, it refers to  
 CC cDNA clones obtained by an oligo-capping method, where none of these  
 CC clones are identical to any known human mRNAs. The present invention  
 CC describes an immunoassay to identify agonists and antagonists, as well as  
 CC antibodies, antisense molecules and siRNAs that can all be used to bind  
 CC to and modulate expression of the cDNA molecules. As such, these  
 CC molecules are useful for diagnostic markers or therapeutic targets for  
 CC the various diseases or morbid states. In particular, they are useful in  
 CC gene therapy for treating osteoporosis, neurological disease, Alzheimer's  
 CC disease, Parkinson's disease, dementia, short memory and various cancers,  
 CC as well as for maintaining equilibrium of sense or motor function, and  
 CC for treating emotional reaction, fear response and panic. Accordingly,  
 CC they exhibit osteopathic, neuroprotective, nootropic, antiparkinsonian,  
 CC cytostatic and tranquiliser activities. This polypeptide is a protein  
 CC encoded by a full length human cDNA sequence of the invention. NOTE: This  
 CC sequence is not given in the sequence listing of the specification but  
 CC can be obtained on CD-ROM from the European Patent Office, Vienna Sub-  
 CC office.

SQ Sequence 214 AA;

Query Match 100.0%; Score 81; DB 8; Length 214;  
 Best Local Similarity 100.0%; Pred. No. 5.9e-06;  
 Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GGTMYPGIADRMQK 15  
 |||||  
 Db 157 GGTMYPGIADRMQK 171

RESULT 15  
 ABG00383  
 ID ABG00383 standard; protein; 258 AA.

XX AC ABG00383;  
 XX DT 13-FEB-2002 (first entry)  
 XX DE Novel human diagnostic protein #374.

XX KW Human; chromosome mapping; gene mapping; gene therapy; forensic;  
 KW food supplement; medical imaging; diagnostic; genetic disorder.  
 XX OS Homo sapiens.

XX PN WO200175067-A2.  
 XX PD 11-OCT-2001.

XX PF 30-MAR-2001; 2001WO-US008631.  
 XX PR 31-MAR-2000; 2000US-00540217.  
 XX PR 23-AUG-2000; 2000US-00649167.

XX PA (HYSE-) HYSEQ INC.

XX PI Drmanac RT, Liu C, Tang YT;  
 XX WPI; 2001-639362/73.  
 DR N-PSDB; AAS64570.

XX New isolated polynucleotide and encoded polypeptides, useful in  
 PT diagnostics, forensics, gene mapping, identification of mutations  
 PT responsible for genetic disorders or other traits and to assess  
 PT biodiversity.

XX PS Claim 20; SEQ ID NO 30742; 103pp; English.

XX The invention relates to isolated polynucleotide (I) and polypeptide (II)  
 CC sequences. (I) is useful as hybridisation probes, polymerase chain  
 CC reaction (PCR) primers, oligomers, and for chromosome and gene mapping,  
 CC and in recombinant production of (II). The polynucleotides are also used  
 CC in diagnostics as expressed sequence tags for identifying expressed  
 CC genes. (I) is useful in gene therapy techniques to restore normal  
 CC activity of (II) or to treat disease states involving (II). (II) is  
 CC useful for generating antibodies against it, detecting or quantitating a  
 CC polypeptide in tissue, as molecular weight markers and as a food  
 CC supplement. (II) and its binding partners are useful in medical imaging  
 CC of sites expressing (II). (I) and (II) are useful for treating disorders  
 CC involving aberrant protein expression or biological activity. The  
 CC polypeptide and polynucleotide sequences have applications in  
 CC diagnostics, forensics, gene mapping, identification of mutations  
 CC responsible for genetic disorders or other traits to assess biodiversity  
 CC and to produce other types of data and products dependent on DNA and  
 CC amino acid sequences. ABG00010-ABG03077 represent novel human diagnostic  
 CC amino acid sequences of the invention. Note: The sequence data for this  
 CC patent did not appear in the printed specification, but was obtained in  
 CC electronic format directly from WIPO at  
 CC ftp.wipo.int/pub/published\_pct\_sequences

XX SQ Sequence 258 AA;

Query Match 100.0%; Score 81; DB 4; Length 258;  
 Best Local Similarity 100.0%; Pred. No. 7.2e-06;  
 Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GGTMYPGIADRMQK 15

Db 192 GGTMYPGIADRMQK 206  
|||||

Search completed: April 8, 2005, 10:50:43  
Job time : 63.4 secs

**THIS PAGE BLANK (USPTO)**

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: April 8, 2005, 09:58:57 ; Search time 17.5333 Seconds  
(without alignments)  
63.863 Million cell updates/sec

Title: US-09-423-351C-4  
Perfect score: 81  
Sequence: 1 GGTMYPGIADRMOK 15

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 513545 seqs, 74649064 residues

Total number of hits satisfying chosen parameters: 513545

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : Issued Patents AA:\*  
1: /cgn2\_6/ptodata/1/iaa/5A\_COMB.pep.\*  
2: /cgn2\_6/ptodata/1/iaa/5B\_COMB.pep.\*  
3: /cgn2\_6/ptodata/1/iaa/6A\_COMB.pep.\*  
4: /cgn2\_6/ptodata/1/iaa/6B\_COMB.pep.\*  
5: /cgn2\_6/ptodata/1/iaa/PCRTUS\_COMB.pep.\*  
6: /cgn2\_6/ptodata/1/iaa/backfiles1.pep.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	81	100.0	60	3	US-09-306-446C-6
2	81	100.0	239	4	US-09-949-016-10409
3	81	100.0	374	3	US-08-609-236-6
4	81	100.0	374	3	US-09-306-446C-2
5	81	100.0	375	3	US-09-106-217-16
6	81	100.0	375	4	US-09-976-594-731
7	81	100.0	376	1	US-08-588-113-2
8	81	100.0	376	4	US-09-949-016-6100
9	81	100.0	377	3	US-09-106-217-2
10	81	100.0	377	4	US-09-919-172-33
11	81	100.0	377	4	US-09-917-254-53
12	81	100.0	386	4	US-09-949-016-7721
13	81	100.0	399	4	US-09-949-016-9424
14	81	100.0	402	4	US-09-949-016-10757
15	81	100.0	404	4	US-09-949-016-11313
16	78	96.3	375	3	US-09-171-337A-7
17	78	96.3	375	3	US-09-631-022-7
18	76	93.8	375	3	US-09-171-337A-8
19	76	93.8	375	3	US-09-631-022-8
20	73	90.1	375	2	US-08-494-151-14
21	73	90.1	377	4	US-09-248-796A-14109
22	66	81.5	21	3	US-08-936-477-1
23	66	81.5	21	4	US-09-786-066-16
24	66	81.5	21	5	PCT-US94-10051-1
25	55	67.9	21	4	US-09-786-066-2
26	55	67.9	21	4	US-09-786-066-9
27	55	67.9	394	4	US-09-949-016-6655

28	55	67.9	406	4	US-09-949-016-7396	Sequence 7396, Ap
29	54	66.7	129	4	US-09-248-796A-14107	Sequence 14107, A
30	43	53.1	471	4	US-09-583-110-3667	Sequence 3667, Ap
31	43	53.1	471	4	US-08-378-213-2	Sequence 2, Appli
32	43	53.1	471	4	US-08-378-213-4	Sequence 4, Appli
33	43	53.1	471	4	US-08-378-213-5	Sequence 5, Appli
34	43	53.1	471	4	US-09-120-044-3	Sequence 3, Appli
35	43	53.1	471	4	US-09-120-044-4	Sequence 4, Appli
36	43	53.1	472	4	US-09-107-433-3566	Sequence 3566, Ap
37	43	53.1	533	4	US-09-902-540-14623	Sequence 14623, A
38	42	51.9	347	4	US-09-543-681A-8255	Sequence 8255, Ap
39	41	50.6	340	4	US-09-214-631-3	Sequence 3, Appli
40	41	50.6	340	4	US-09-051-994-2	Sequence 2, Appli
41	41	50.6	340	4	US-08-635-130A-4	Sequence 4, Appli
42	41	50.6	340	4	US-09-949-016-6076	Sequence 6076, Ap
43	41	50.6	362	4	US-09-949-016-7725	Sequence 7725, Ap
44	41	50.6	376	4	US-09-538-092-1109	Sequence 1109, Ap
45	41	50.6	376	4	US-09-538-092-1110	Sequence 1110, Ap

## ALIGNMENTS

RESULT 1  
US-09-306-446C-6  
; Sequence 6, Application US/09306446C  
; Patent No. 6372959  
; GENERAL INFORMATION:  
; APPLICANT: KIM, Dong Soo  
; APPLICANT: KIM, Chul Geun  
; APPLICANT: NAM, Yoon Kwon  
; APPLICANT: NOH, Jae Koo  
; APPLICANT: CHO, Kyoun Nam  
; TITLE OF INVENTION: EXPRESSION VECTOR OF MUD LOACH GROWTH HORMONE GENE  
; FILE REFERENCE: P06344050/BAS  
; CURRENT APPLICATION NUMBER: US/09/306.446C  
; CURRENT FILING DATE: 1999-05-06  
; PRIOR APPLICATION NUMBER: KR 98/20255  
; PRIOR FILING DATE: 1998-06-01  
; NUMBER OF SEQ ID NOS: 33  
; SOFTWARE: Patent in Ver. 2.0  
; SEQ ID NO 6  
; LENGTH: 60  
; TYPE: PRT  
; ORGANISM: Misgurnus mizolepus  
US-09-306-446C-6

Query Match 100.0%; Score 81; DB 3; Length 60;  
Best Local Similarity 100.0%; Pred. No. 5.9e-07;  
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGTMYPGIADRMOK 15  
DB 33 GGTMYPGIADRMOK 47

RESULT 2  
US-09-949-016-10409  
; Sequence 10409, Application US/09949016  
; Patent No. 6812339  
; GENERAL INFORMATION:  
; APPLICANT: VENTER, J. Craig et al.  
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF  
; FILE REFERENCE: CLO001307  
; CURRENT APPLICATION NUMBER: US/09/949.016  
; CURRENT FILING DATE: 2000-04-14  
; PRIOR APPLICATION NUMBER: 60/241,755  
; PRIOR FILING DATE: 2000-10-20  
; PRIOR APPLICATION NUMBER: 60/237,768  
; PRIOR FILING DATE: 2000-10-03  
; PRIOR APPLICATION NUMBER: 60/231,498  
; PRIOR FILING DATE: 2000-09-08

; NUMBER OF SEQ ID NOS: 207012  
; SOFTWARE: FastSEQ for Windows Version 4.0  
; SEQ ID NO 10409  
; LENGTH: 239  
; TYPE: PRT  
; ORGANISM: Human  
US-09-949-016-10409

Query Match 100.0%; Score 81; DB 4; Length 239;  
Best Local Similarity 100.0%; Pred. No. 2.7e-06;  
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGTMYPGIADRMQK 15  
|||||  
DB 165 GGTMYPGIADRMQK 179

RESULT 3  
US-08-609-236-6  
; Sequence 6, Application US/08609236  
; Patent No. 6087398  
; GENERAL INFORMATION:  
; APPLICANT: Steven R. Goodman  
; TITLE OF INVENTION: No. 6087398el Sickel Cell Anemia Treatment  
; NUMBER OF SEQUENCES: 6  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: McGregor & Adler, LLP  
; STREET: 8011 Candle Lane  
; CITY: Houston  
; STATE: Texas  
; COUNTRY: USA  
; ZIP: 77071  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: Apple Macintosh  
; OPERATING SYSTEM: Macintosh  
; SOFTWARE: Microsoft Word for Macintosh  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/609,236  
; FILING DATE: March 1, 1996  
; CLASSIFICATION: 514  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 60/002,288  
; FILING DATE: August 14, 1995  
; CLASSIFICATION: 514  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Benjamin Aaron Adler, Ph.D.  
; REGISTRATION NUMBER: 35,423  
; REFERENCE/DOCKET NUMBER: D5807  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 713-777-2321  
; TELEFAX: 713-777-6908  
; INFORMATION FOR SEQ ID NO: 6:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 374  
; TYPE: Amino acid  
; STRANDEDNESS:  
; TOPOLOGY: Linear  
; MOLECULE TYPE: Protein  
; DESCRIPTION: Protein  
; HYPOTHETICAL: NO  
; ANTI-SENSE: NO  
; FRAGMENT TYPE:  
; ORIGINAL SOURCE:  
; STRAIN:  
; INDIVIDUAL ISOLATE:  
; DEVELOPMENTAL STAGE:  
; TISSUE TYPE:  
; CELL TYPE:  
; CELL LINE:

US-08-609-236-6 Query Match 100.0%; Score 81; DB 3; Length 374;

Best Local Similarity 100.0%; Pred. No. 4.4e-06;  
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGTMYPGIADRMQK 15  
|||||  
DB 300 GGTMYPGIADRMQK 314

RESULT 4  
US-09-306-446C-2  
; Sequence 2, Application US/09306446C  
; Patent No. 6372959  
; GENERAL INFORMATION:  
; APPLICANT: KIM, Dong Soo  
; APPLICANT: KIM, Chul Geun  
; APPLICANT: NAM, Yoon Kwon  
; APPLICANT: NOH, Jae Koo  
; APPLICANT: CHO, Kyou Nam  
; TITLE OF INVENTION: EXPRESSION VECTOR OF MUD LOACH GROWTH HORMONE GENE  
; FILE REFERENCE: P06344US0/BAS  
; CURRENT APPLICATION NUMBER: US/09/306,446C  
; CURRENT FILING DATE: 1999-05-06  
; PRIOR APPLICATION NUMBER: KR 98/20255  
; PRIOR FILING DATE: 1998-06-01  
; NUMBER OF SEQ ID NOS: 33  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 2  
; LENGTH: 374  
; TYPE: PRT  
; ORGANISM: Misgurnus mizolepus  
US-09-306-446C-2

Query Match 100.0%; Score 81; DB 3; Length 374;  
Best Local Similarity 100.0%; Pred. No. 4.4e-06;  
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGTMYPGIADRMQK 15  
|||||  
DB 300 GGTMYPGIADRMQK 314

RESULT 5  
US-09-106-217-16  
; Sequence 16, Application US/09106217  
; Patent No. 6063576  
; GENERAL INFORMATION:  
; APPLICANT: Keating, Mark T.  
; APPLICANT: Olson, Timothy M.  
; TITLE OF INVENTION: Actin Mutations in Dilated  
; TITLE OF INVENTION: Cardiomyopathy, a Heritable Form of Heart Failure  
; NUMBER OF SEQUENCES: 18  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Rothwell, Figg, Ernst & Kurz, P.C.  
; STREET: 555 Thirteenth Street, N.W., Suite 701 East  
; STREET: Tower  
; CITY: Washington  
; STATE: DC  
; COUNTRY: U.S.A.  
; ZIP: 20004  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/09/106,217  
; FILING DATE:  
; CLASSIFICATION:  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Saxe, Stephen A.  
; REGISTRATION NUMBER: 38,609  
; REFERENCE/DOCKET NUMBER: 2323-125  
; TELECOMMUNICATION INFORMATION:

TELEPHONE: 202-783-6040  
TELEFAX: 202-783-6031  
INFORMATION FOR SEQ ID NO: 16:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 375 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-09-106-217-16

Query Match 100.0%; Score 81; DB 3; Length 375;  
Best Local Similarity 100.0%; Pred. No. 4.4e-06;  
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGTWYPGIADRMQK 15  
|||||  
DB 301 GGTWYPGIADRMQK 315

## RESULT 6

US-09-594-594-731  
Sequence 731, Application US/09976594  
Patent No. 6673549

## GENERAL INFORMATION:

APPLICANT: Furness, Michael  
APPLICANT: Buchbinder, Jenny  
TITLE OF INVENTION: GENES EXPRESSED IN C3A LIVER CELL CULTURES TREATED WITH STEROIDS  
FILE REFERENCE: PA-0041 US  
CURRENT APPLICATION NUMBER: US/09/976,594  
CURRENT FILING DATE: 2001-10-12  
PRIOR APPLICATION NUMBER: 60/240,409  
PRIOR FILING DATE: 2000-10-12  
NUMBER OF SEQ ID NOS: 1143  
SOFTWARE: PERL Program  
SEQ ID NO: 731  
LENGTH: 375  
TYPE: PRT  
ORGANISM: Homo sapiens  
NAME/KEY: misc\_feature  
OTHER INFORMATION: Incyte ID No. 6673549 1837317CDD1  
US-09-594-594-731

Query Match 100.0%; Score 81; DB 4; Length 375;  
Best Local Similarity 100.0%; Pred. No. 4.4e-06;  
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGTWYPGIADRMQK 15  
|||||  
DB 301 GGTWYPGIADRMQK 315

## RESULT 7

US-08-588-113-2  
Sequence 2, Application US/08588113  
Patent No. 5710003

## GENERAL INFORMATION:

APPLICANT: McHugh, Kirk M.  
TITLE OF INVENTION: DIAGNOSTIC TEST FOR DETERMINING  
TITLE OF INVENTION: MALIGNANCY OF SMOOTH MUSCLE TUMORS  
NUMBER OF SEQUENCES: 4  
CORRESPONDENCE ADDRESS:

ADDRESSEE: Woodcock Washburn Kurtz Mackiewicz &  
ADDRESSEE: No. 571003r18  
STREET: One Liberty Place - 46th Floor  
CITY: Philadelphia  
STATE: PA  
COUNTRY: USA  
ZIP: 19103

COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: PatentIn Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/588,113  
FILING DATE:

## CLASSIFICATION: 436

ATTORNEY/AGENT INFORMATION:  
NAME: Ralph, Rebecca L.  
REGISTRATION NUMBER: 35,152  
REFERENCE/DOCKET NUMBER: TJU-1652  
TELECOMMUNICATION INFORMATION:

TELEPHONE: 215-568-3100

TELEFAX: 215-568-3439

INFORMATION FOR SEQ ID NO: 2:

SEQUENCE CHARACTERISTICS:

LENGTH: 376 amino acids

TYPE: amino acid

TOPOLOGY: linear

MOLECULE TYPE: protein

US-08-588-113-2

Query Match 100.0%; Score 81; DB 1; Length 376;  
Best Local Similarity 100.0%; Pred. No. 4.4e-06;  
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGTWYPGIADRMQK 15  
|||||  
DB 302 GGTWYPGIADRMQK 316

## RESULT 8

US-09-949-016-6100  
Sequence 6100, Application US/09949016  
Patent No. 6812339

## GENERAL INFORMATION:

APPLICANT: VENTER, J. Craig et al.  
TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED  
WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF  
FILE REFERENCE: CL001307  
CURRENT APPLICATION NUMBER: US/09/949,016

CURRENT FILING DATE: 2000-04-14

PRIOR APPLICATION NUMBER: 60/241,755

PRIOR FILING DATE: 2000-10-20

PRIOR APPLICATION NUMBER: 60/237,768

PRIOR FILING DATE: 2000-10-03

PRIOR APPLICATION NUMBER: 60/231,498

PRIOR FILING DATE: 2000-09-08

NUMBER OF SEQ ID NOS: 207012

SOFTWARE: FastSeq for Windows Version 4.0

SEQ ID NO: 6100

LENGTH: 376

TYPE: PRT

ORGANISM: Human

US-09-949-016-6100

Query Match 100.0%; Score 81; DB 4; Length 376;  
Best Local Similarity 100.0%; Pred. No. 4.4e-06;  
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGTWYPGIADRMQK 15  
|||||  
DB 302 GGTWYPGIADRMQK 316

## RESULT 9

US-09-106-217-2  
Sequence 2, Application US/09106217  
Patent No. 6063576

## GENERAL INFORMATION:

APPLICANT: Keating, Mark T.

APPLICANT: Olson, Timothy M.

TITLE OF INVENTION: Actin Mutations in Dilated

TITLE OF INVENTION: Cardiomyopathy, a Heritable Form of Heart Failure

NUMBER OF SEQUENCES: 18

;; CORRESPONDENCE ADDRESS:  
;; ADDRESSEE: Rothwell, Figg, Ernst & Kurz, P.C.  
;; STREET: 555 Thirteenth Street, N.W., Suite 701 East  
;; STREET: Tower  
;; CITY: Washington  
;; STATE: DC  
;; COUNTRY: U.S.A.  
;; ZIP: 20004  
;; COMPUTER READABLE FORM:  
;; MEDIUM TYPE: Floppy disk  
;; COMPUTER: IBM PC compatible  
;; OPERATING SYSTEM: PC-DOS/MS-DOS  
;; SOFTWARE: Patent In Release #1.0, Version #1.30  
;; CURRENT APPLICATION DATA:  
;; APPLICATION NUMBER: US/09/106,217  
;; FILING DATE:  
;; CLASSIFICATION:  
;; ATTORNEY/AGENT INFORMATION:  
;; NAME: Saxe, Stephen A.  
;; REGISTRATION NUMBER: 38,609  
;; REFERENCE/DOCKET NUMBER: 2323-125  
;; TELECOMMUNICATION INFORMATION:  
;; TELEPHONE: 202-783-6040  
;; TELEFAX: 202-783-6031  
;; INFORMATION FOR SEQ ID NO: 2:  
;; SEQUENCE CHARACTERISTICS:  
;; LENGTH: 377 amino acids  
;; TYPE: amino acid  
;; TOPOLOGY: linear  
;; MOLECULE TYPE: protein  
US-09-106-217-2

Query Match 100.0%; Score 81; DB 3; Length 377;  
Best Local Similarity 100.0%; Pred. No. 4.4e-06;  
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGTMYPGIADRMQK 15  
|||  
Db 303 GGTMYPGIADRMQK 317

RESULT 10  
US-09-919-172-33  
; Sequence 33, Application US/09919172  
; Patent No. 6673545  
; GENERAL INFORMATION:  
; APPLICANT: Paris, Mary  
; TITLE OF INVENTION: Turner, Christopher M.  
; FILE REFERENCE: PA-0036 US  
; CURRENT APPLICATION NUMBER: US/09/919,172  
; CURRENT FILING DATE: 2001-07-30  
; PRIOR APPLICATION NUMBER: 60/222,469  
; PRIOR FILING DATE: 2000-07-28  
; NUMBER OF SEQ ID NOS: 102  
; SOFTWARE: PERL Program  
; SEQ ID NO 33  
; LENGTH: 377  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
; FEATURE:  
; NAME/KEY: misc feature  
; OTHER INFORMATION: Incyte ID No. 6673545 1709118CD1  
US-09-919-172-33

Query Match 100.0%; Score 81; DB 4; Length 377;  
Best Local Similarity 100.0%; Pred. No. 4.4e-06;  
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGTMYPGIADRMQK 15  
|||  
Db 303 GGTMYPGIADRMQK 317

RESULT 11  
US-09-917-254-53  
; Sequence 53, Application US/09917254  
; Patent No. 6703204  
; GENERAL INFORMATION:  
; APPLICANT: Mutter, George  
; APPLICANT: Baak, Jan  
; TITLE OF INVENTION: Prognostic Classification of Breast Cancer  
; FILE REFERENCE: B0801/7224(JRV)  
; CURRENT APPLICATION NUMBER: US/09/917,254  
; CURRENT FILING DATE: 2001-07-27  
; PRIOR APPLICATION NUMBER: US 60/222,093  
; PRIOR FILING DATE: 2000-07-28  
; NUMBER OF SEQ ID NOS: 102  
; SOFTWARE: Patent In version 3.0  
; SEQ ID NO 53  
; LENGTH: 377  
; TYPE: PRT  
; ORGANISM: Homo Sapiens  
US-09-917-254-53

Query Match 100.0%; Score 81; DB 4; Length 377;  
Best Local Similarity 100.0%; Pred. No. 4.4e-06;  
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGTMYPGIADRMQK 15  
|||  
Db 303 GGTMYPGIADRMQK 317

## RESULT 12

US-09-949-016-7721  
; Sequence 7721, Application US/09949016  
; Patent No. 6812339  
; GENERAL INFORMATION:  
; APPLICANT: VENTER, J. Craig et al.  
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED  
; FILE REFERENCE: CLO01307  
; CURRENT APPLICATION NUMBER: US/09/949,016  
; CURRENT FILING DATE: 2000-04-14  
; PRIOR APPLICATION NUMBER: 60/241,755  
; PRIOR FILING DATE: 2000-10-20  
; PRIOR APPLICATION NUMBER: 60/237,768  
; PRIOR FILING DATE: 2000-10-03  
; PRIOR APPLICATION NUMBER: 60/231,498  
; PRIOR FILING DATE: 2000-09-08  
; NUMBER OF SEQ ID NOS: 207012  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 7721  
; LENGTH: 386  
; TYPE: PRT  
; ORGANISM: Human  
US-09-949-016-7721

Query Match 100.0%; Score 81; DB 4; Length 386;  
Best Local Similarity 100.0%; Pred. No. 4.5e-06;  
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGTMYPGIADRMQK 15  
|||  
Db 312 GGTMYPGIADRMQK 326

## RESULT 13

US-09-949-016-9424  
; Sequence 9424, Application US/09949016  
; Patent No. 6812339  
; GENERAL INFORMATION:  
; APPLICANT: VENTER, J. Craig et al.  
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED  
; FILE REFERENCE: CLO01307  
; CURRENT APPLICATION NUMBER: US/09/949,016  
; CURRENT FILING DATE: 2000-04-14  
; PRIOR APPLICATION NUMBER: 60/241,755  
; PRIOR FILING DATE: 2000-10-20  
; PRIOR APPLICATION NUMBER: 60/237,768  
; PRIOR FILING DATE: 2000-10-03  
; PRIOR APPLICATION NUMBER: 60/231,498  
; PRIOR FILING DATE: 2000-09-08  
; NUMBER OF SEQ ID NOS: 207012  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 7721  
; LENGTH: 386  
; TYPE: PRT  
; ORGANISM: Human  
US-09-949-016-9424

Query Match 100.0%; Score 81; DB 4; Length 386;  
Best Local Similarity 100.0%; Pred. No. 4.5e-06;  
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGTMYPGIADRMQK 15  
|||  
Db 312 GGTMYPGIADRMQK 326



; FILE REFERENCE: CL001307  
; CURRENT APPLICATION NUMBER: US/09/949,016  
; PRIOR FILING DATE: 2000-04-14  
; PRIOR APPLICATION NUMBER: 60/241,755  
; PRIOR FILING DATE: 2000-10-20  
; PRIOR APPLICATION NUMBER: 60/237,768  
; PRIOR FILING DATE: 2000-10-03  
; PRIOR APPLICATION NUMBER: 60/231,498  
; PRIOR FILING DATE: 2000-09-08  
; NUMBER OF SEQ ID NOS: 207012  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 9424  
; LENGTH: 399  
; TYPE: PRT  
; ORGANISM: Human  
US-09-949-016-9424

Query Match 100.0%; Score 81; DB 4; Length 399;  
Best Local Similarity 100.0%; Pred. No. 4.7e-06;  
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGTMYPGIADRMQK 15  
Db 325 GGTMYPGIADRMQK 339

## RESULT 14

US-09-949-016-10757  
; Sequence 10757, Application US/09949016  
; Patent No. 6812339  
; GENERAL INFORMATION:  
; APPLICANT: VENTER, J. Craig et al.  
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED  
; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF  
; FILE REFERENCE: CL001307  
; CURRENT APPLICATION NUMBER: US/09/949,016  
; PRIOR FILING DATE: 2000-04-14  
; PRIOR APPLICATION NUMBER: 60/241,755  
; PRIOR FILING DATE: 2000-10-20  
; PRIOR APPLICATION NUMBER: 60/237,768  
; PRIOR FILING DATE: 2000-10-03  
; PRIOR APPLICATION NUMBER: 60/231,498  
; PRIOR FILING DATE: 2000-09-08  
; NUMBER OF SEQ ID NOS: 207012  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 10757  
; LENGTH: 402  
; TYPE: PRT  
; ORGANISM: Human  
US-09-949-016-10757

Query Match 100.0%; Score 81; DB 4; Length 402;  
Best Local Similarity 100.0%; Pred. No. 4.7e-06;  
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGTMYPGIADRMQK 15  
Db 328 GGTMYPGIADRMQK 342

## RESULT 15

US-09-949-016-11313  
; Sequence 11313, Application US/09949016  
; Patent No. 6812339  
; GENERAL INFORMATION:  
; APPLICANT: VENTER, J. Craig et al.  
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED  
; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF  
; FILE REFERENCE: CL001307  
; CURRENT APPLICATION NUMBER: US/09/949,016  
; PRIOR FILING DATE: 2000-04-14  
; PRIOR APPLICATION NUMBER: 60/241,755  
; PRIOR FILING DATE: 2000-10-20

; FILE REFERENCE: CL001307  
; CURRENT APPLICATION NUMBER: 60/237,768  
; PRIOR FILING DATE: 2000-10-03  
; PRIOR APPLICATION NUMBER: 60/231,498  
; PRIOR FILING DATE: 2000-09-08  
; NUMBER OF SEQ ID NOS: 207012  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 11313  
; LENGTH: 404  
; TYPE: PRT  
; ORGANISM: Human  
US-09-949-016-11313

Query Match 100.0%; Score 81; DB 4; Length 404;  
Best Local Similarity 100.0%; Pred. No. 4.7e-06;  
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGTMYPGIADRMQK 15  
Db 330 GGTMYPGIADRMQK 344

Search completed: April 8, 2005, 12:07:35  
Job time : 18.5333 secs

**THIS PAGE BLANK (USPTO)**

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: April 8, 2005, 10:53:18 ; Search time 42.6667 Seconds  
(without alignments)  
116.718 Million cell updates/sec

**Title:** US-09-423-351C-4

Perfect score:

Sequence: 1 GGTMYPGIADRMQK 15

Scoring table: BLOSUM62

Scoring cards: [Download](#)  
Gapop 10.0 , Gapext 0.5

Searched: 1418010 sears. 331997259 residues

Total number of hits satisfying chosen parameters: 1418010

Minimum DB seq length: 0

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Fast-Processing: Minimum Match 0%  
Maximum Match 100%

Listing first 45 summaries

```
Database : Published Applications_AA.*
1: /cgn2_6/prodata/1/pubpaa/us07_PUBCOMB.pcp.*
2: /cgn2_6/prodata/1/pubpaa/PCT_NEW_PUB.pcp.*
3: /cgn2_6/prodata/1/pubpaa/US06_NEW_PUB.pcp.*
4: /cgn2_6/prodata/1/pubpaa/US06_PUBCOMB.pcp.*
5: /cgn2_6/prodata/1/pubpaa/US07_NEW_PUB.pcp.*
6: /cgn2_6/prodata/1/pubpaa/PCTU5_PUBCOMB.pcp.*
7: /cgn2_6/prodata/1/pubpaa/US08_NEW_PUB.pcp.*
8: /cgn2_6/prodata/1/pubpaa/US08_PUBCOMB.pcp.*
9: /cgn2_6/prodata/1/pubpaa/US09A_PUBCOMB.pcp.*
10: /cgn2_6/prodata/1/pubpaa/US09B_PUBCOMB.pcp.*
11: /cgn2_6/prodata/1/pubpaa/US09C_PUBCOMB.pcp.*
12: /cgn2_6/prodata/1/pubpaa/US09_NEW_PUB.pcp.*
13: /cgn2_6/prodata/1/pubpaa/US10A_PUBCOMB.pcp.*
14: /cgn2_6/prodata/1/pubpaa/US10B_PUBCOMB.pcp.*
15: /cgn2_6/prodata/1/pubpaa/US10C_PUBCOMB.pcp.*
16: /cgn2_6/prodata/1/pubpaa/US10D_PUBCOMB.pcp.*
17: /cgn2_6/prodata/1/pubpaa/US11_NEW_PUB.pcp.*
18: /cgn2_6/prodata/1/pubpaa/US11_NEW_PUB.pcp.*
19: /cgn2_6/prodata/1/pubpaa/US60_NEW_PUB.pcp.*
20: /cgn2_6/prodata/1/pubpaa/US60_PUBCOMB.pcp.*
```

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	81	100.0	100	9	US-09-796-692-1768	Sequence 1768, Ap
2	81	100.0	100	9	US-09-796-692-2271	Sequence 2271, Ap
3	81	100.0	100	14	US-10-040-862-1768	Sequence 1768, Ap
4	81	100.0	100	14	US-10-040-862-2271	Sequence 2271, Ap
5	81	100.0	100	15	US-10-057-475B-1768	Sequence 1768, Ap
6	81	100.0	100	15	US-10-057-475B-2271	Sequence 2271, Ap
7	81	100.0	100	15	US-10-154-884B-1768	Sequence 1768, Ap
8	81	100.0	100	15	US-10-154-884B-2271	Sequence 2271, Ap
9	81	100.0	100	16	US-10-764-324-1768	Sequence 1768, Ap
10	81	100.0	100	16	US-10-764-324-2271	Sequence 2271, Ap
11	81	100.0	110	15	US-10-264-049-4060	Sequence 4060, Ap
12	81	100.0	133	15	US-10-424-599-143033	Sequence 143033, Ap
13	81	100.0	145	15	US-10-264-049-4259	Sequence 4259, Ap

; PRIOR FILING DATE: 2000-08-07  
 ; NUMBER OF SEQ ID NOS: 9597  
 ; SOFTWARE: FastSeq for Windows Version 3.0  
 ; SEQ ID NO 1768  
 ; LENGTH: 100  
 ; TYPE: PRT  
 ; ORGANISM: Homo sapiens  
 US-09-796-692-1768

Query Match 100.0%; Score 81; DB 9; Length 100;  
 Best Local Similarity 100.0%; Pred. No. 2.6e-06;  
 Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGTMYPGIADRMQK 15  
 Db 46 GGTMYPGIADRMQK 60

RESULT 2

US-09-796-692-2271  
 ; Sequence 2271, Application US/09796692  
 ; Publication No. US20020198362A1  
 ; GENERAL INFORMATION:

; APPLICANT: Gaiger, Alexander  
 ; APPLICANT: Algate, Paul A.  
 ; APPLICANT: Mannion, Jane  
 ; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE DETECTION, DIAGNOSIS AND THERAPY  
 ; FILE REFERENCE: 2077.001200  
 ; CURRENT APPLICATION NUMBER: US/09/796,692  
 ; PRIOR FILING DATE: 2001-03-01  
 ; PRIOR APPLICATION NUMBER: 60/186,126  
 ; PRIOR FILING DATE: 2000-03-01  
 ; PRIOR APPLICATION NUMBER: 60/190,479  
 ; PRIOR FILING DATE: 2000-03-17  
 ; PRIOR APPLICATION NUMBER: 60/200,545  
 ; PRIOR FILING DATE: 2000-04-27  
 ; PRIOR APPLICATION NUMBER: 60/200,303  
 ; PRIOR FILING DATE: 2000-04-28  
 ; PRIOR APPLICATION NUMBER: 60/200,779  
 ; PRIOR FILING DATE: 2000-04-28  
 ; PRIOR APPLICATION NUMBER: 60/200,999  
 ; PRIOR FILING DATE: 2000-05-01  
 ; PRIOR APPLICATION NUMBER: 60/202,084  
 ; PRIOR FILING DATE: 2000-05-04  
 ; PRIOR APPLICATION NUMBER: 60/206,201  
 ; PRIOR FILING DATE: 2000-05-22  
 ; PRIOR APPLICATION NUMBER: 60/218,950  
 ; PRIOR FILING DATE: 2000-07-14  
 ; PRIOR APPLICATION NUMBER: 60/222,903  
 ; PRIOR FILING DATE: 2000-08-03  
 ; PRIOR APPLICATION NUMBER: 60/223,416  
 ; PRIOR FILING DATE: 2000-08-04  
 ; PRIOR APPLICATION NUMBER: 60/223,378  
 ; PRIOR FILING DATE: 2000-08-07  
 ; NUMBER OF SEQ ID NOS: 9597  
 ; SOFTWARE: FastSeq for Windows Version 3.0  
 ; SEQ ID NO 2271  
 ; LENGTH: 100  
 ; TYPE: PRT  
 ; ORGANISM: Homo sapiens  
 US-09-796-692-2271

Query Match 100.0%; Score 81; DB 9; Length 100;  
 Best Local Similarity 100.0%; Pred. No. 2.6e-06;  
 Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGTMYPGIADRMQK 15  
 Db 46 GGTMYPGIADRMQK 60

RESULT 3

US-10-040-862-1768  
 ; Sequence 1768, Application US/10040862  
 ; Publication No. US20030078396A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Gaiger, Alexander  
 ; APPLICANT: Algate, Paul A.  
 ; APPLICANT: Mannion, Jane  
 ; APPLICANT: Retter, Marc  
 ; APPLICANT: Corixa Corporation  
 ; TITLE OF INVENTION: Compositions and Methods for the Detection, Diagnosis and Therapy  
 ; FILE REFERENCE: 014058-013520US  
 ; CURRENT APPLICATION NUMBER: US/10/040,862  
 ; CURRENT FILING DATE: 2001-11-06  
 ; PRIOR APPLICATION NUMBER: US 60/186,126  
 ; PRIOR FILING DATE: 2000-03-01  
 ; PRIOR APPLICATION NUMBER: US 60/190,479  
 ; PRIOR FILING DATE: 2000-03-17  
 ; PRIOR APPLICATION NUMBER: US 60/200,545  
 ; PRIOR FILING DATE: 2000-04-27  
 ; PRIOR APPLICATION NUMBER: US 60/200,303  
 ; PRIOR FILING DATE: 2000-04-28  
 ; PRIOR APPLICATION NUMBER: US 60/200,779  
 ; PRIOR FILING DATE: 2000-04-28  
 ; PRIOR APPLICATION NUMBER: US 60/200,999  
 ; PRIOR FILING DATE: 2000-05-01  
 ; PRIOR APPLICATION NUMBER: US 60/202,084  
 ; PRIOR FILING DATE: 2000-05-04  
 ; PRIOR APPLICATION NUMBER: US 60/206,201  
 ; PRIOR FILING DATE: 2000-05-22  
 ; PRIOR APPLICATION NUMBER: US 60/218,950  
 ; PRIOR FILING DATE: 2000-07-14  
 ; PRIOR APPLICATION NUMBER: US 60/222,903  
 ; PRIOR FILING DATE: 2000-08-03  
 ; PRIOR APPLICATION NUMBER: US 60/223,416  
 ; PRIOR FILING DATE: 2000-08-04  
 ; PRIOR APPLICATION NUMBER: US 60/223,378  
 ; PRIOR FILING DATE: 2000-08-07  
 ; PRIOR APPLICATION NUMBER: US 09/796,692  
 ; NUMBER OF SEQ ID NOS: 10467  
 ; SOFTWARE: FastSeq for Windows Version 3.0  
 ; SEQ ID NO 1768  
 ; LENGTH: 100  
 ; TYPE: PRT  
 ; ORGANISM: Homo sapiens  
 US-10-040-862-1768

Query Match 100.0%; Score 81; DB 14; Length 100;  
 Best Local Similarity 100.0%; Pred. No. 2.6e-06;  
 Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGTMYPGIADRMQK 15  
 Db 46 GGTMYPGIADRMQK 60

RESULT 4

US-10-040-862-2271  
 ; Sequence 2271, Application US/10040862  
 ; Publication No. US20030078396A1  
 ; GENERAL INFORMATION:

; APPLICANT: Gaiger, Alexander  
 ; APPLICANT: Algate, Paul A.  
 ; APPLICANT: Mannion, Jane  
 ; APPLICANT: Retter, Marc  
 ; APPLICANT: Corixa Corporation  
 ; TITLE OF INVENTION: Compositions and Methods for the Detection, Diagnosis and Therapy  
 ; FILE REFERENCE: 014058-013520US  
 ; CURRENT APPLICATION NUMBER: US/10/040,862  
 ; CURRENT FILING DATE: 2001-11-06  
 ; PRIOR APPLICATION NUMBER: US 60/186,126

```

; PRIOR FILING DATE: 2000-03-01
; PRIOR APPLICATION NUMBER: US 60/190,479
; PRIOR FILING DATE: 2000-03-17
; PRIOR APPLICATION NUMBER: US 60/200,545
; PRIOR FILING DATE: 2000-04-27
; PRIOR APPLICATION NUMBER: US 60/200,303
; PRIOR FILING DATE: 2000-04-28
; PRIOR APPLICATION NUMBER: US 60/200,779
; PRIOR FILING DATE: 2000-04-28
; PRIOR APPLICATION NUMBER: US 60/200,999
; PRIOR FILING DATE: 2000-05-01
; PRIOR APPLICATION NUMBER: US 60/202,084
; PRIOR FILING DATE: 2000-05-04
; PRIOR APPLICATION NUMBER: US 60/206,201
; PRIOR FILING DATE: 2000-05-22
; PRIOR APPLICATION NUMBER: US 60/218,950
; PRIOR FILING DATE: 2000-07-14
; PRIOR APPLICATION NUMBER: US 60/222,903
; PRIOR FILING DATE: 2000-08-03
; PRIOR APPLICATION NUMBER: US 60/223,416
; PRIOR FILING DATE: 2000-08-04
; PRIOR APPLICATION NUMBER: US 60/223,378
; PRIOR FILING DATE: 2000-08-07
; PRIOR APPLICATION NUMBER: US 09/796,692
; PRIOR FILING DATE: 2001-03-01
; NUMBER OF SEQ ID NOS: 10467
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 2271
; LENGTH: 100
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-040-862-2271

```

```

Query Match 100.0%; Score 81; DB 14; Length 100;
Best Local Similarity 100.0%; Pred. No. 2.6e-06;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

Qy 1 GGTMYPGIADRMQK 15
Db 46 GGTMYPGIADRMQK 60

```

```

RESULT 5
US-10-057-475B-1768
; Sequence 1768, Application US/10057475B
; Publication No. US20040002068A1
; GENERAL INFORMATION:
; APPLICANT: Gaiger, Alexander
; APPLICANT: Algate, Paul A.
; APPLICANT: Mannion, Jane
; APPLICANT: Clapper, Jonathan David
; APPLICANT: Wang, Aijun
; APPLICANT: Ordenez, Nadia
; APPLICANT: Carter, Lauren
; APPLICANT: McNeill, Patricia Dianne
; APPLICANT: Corixa Corporation
; TITLE OF INVENTION: Compositions and Methods for the Detection, Diagnosis and Therapy
; TITLE OF INVENTION: Hematological Malignancies
; FILE REFERENCE: 014058-014402US
; CURRENT APPLICATION NUMBER: US/10/057,475B
; CURRENT FILING DATE: 2002-01-22
; PRIOR FILING DATE: 2000-03-01
; PRIOR APPLICATION NUMBER: US 60/186,126
; PRIOR FILING DATE: 2000-03-17
; PRIOR APPLICATION NUMBER: US 60/190,479
; PRIOR FILING DATE: 2000-04-27
; PRIOR APPLICATION NUMBER: US 60/200,545
; PRIOR FILING DATE: 2000-04-28
; PRIOR APPLICATION NUMBER: US 60/200,303
; PRIOR FILING DATE: 2000-05-01
; PRIOR APPLICATION NUMBER: US 60/202,084
; PRIOR FILING DATE: 2000-05-04
; PRIOR APPLICATION NUMBER: US 60/206,201
; PRIOR FILING DATE: 2000-05-22
; PRIOR APPLICATION NUMBER: US 60/218,950
; PRIOR FILING DATE: 2000-07-14
; PRIOR APPLICATION NUMBER: US 60/222,903
; PRIOR FILING DATE: 2000-08-03
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 10467
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 2271
; LENGTH: 100
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-057-475B-2271

```

```

; PRIOR APPLICATION NUMBER: US 60/202,084
; PRIOR FILING DATE: 2000-05-04
; PRIOR APPLICATION NUMBER: US 60/206,201
; PRIOR FILING DATE: 2000-05-22
; PRIOR APPLICATION NUMBER: US 60/218,950
; PRIOR FILING DATE: 2000-07-14
; PRIOR APPLICATION NUMBER: US 60/222,903
; PRIOR FILING DATE: 2000-08-03
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 10979
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 1768
; LENGTH: 100
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-057-475B-1768

```

```

Query Match 100.0%; Score 81; DB 15; Length 100;
Best Local Similarity 100.0%; Pred. No. 2.6e-06;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

Qy 1 GGTMYPGIADRMQK 15
Db 46 GGTMYPGIADRMQK 60

```

```

RESULT 6
US-10-057-475B-2271
; Sequence 2271, Application US/10057475B
; Publication No. US20040002068A1
; GENERAL INFORMATION:
; APPLICANT: Gaiger, Alexander
; APPLICANT: Algate, Paul A.
; APPLICANT: Mannion, Jane
; APPLICANT: Clapper, Jonathan David
; APPLICANT: Wang, Aijun
; APPLICANT: Ordenez, Nadia
; APPLICANT: Carter, Lauren
; APPLICANT: McNeill, Patricia Dianne
; APPLICANT: Corixa Corporation
; TITLE OF INVENTION: Compositions and Methods for the Detection, Diagnosis and Therapy
; TITLE OF INVENTION: Hematological Malignancies
; FILE REFERENCE: 014058-014402US
; CURRENT APPLICATION NUMBER: US/10/057,475B
; CURRENT FILING DATE: 2002-01-22
; PRIOR FILING DATE: 2000-03-01
; PRIOR APPLICATION NUMBER: US 60/186,126
; PRIOR FILING DATE: 2000-03-17
; PRIOR APPLICATION NUMBER: US 60/190,479
; PRIOR FILING DATE: 2000-04-27
; PRIOR APPLICATION NUMBER: US 60/200,545
; PRIOR FILING DATE: 2000-04-28
; PRIOR APPLICATION NUMBER: US 60/200,303
; PRIOR FILING DATE: 2000-04-28
; PRIOR APPLICATION NUMBER: US 60/200,779
; PRIOR FILING DATE: 2000-04-28
; PRIOR APPLICATION NUMBER: US 60/200,999
; PRIOR FILING DATE: 2000-05-01
; PRIOR APPLICATION NUMBER: US 60/202,084
; PRIOR FILING DATE: 2000-05-04
; PRIOR APPLICATION NUMBER: US 60/206,201
; PRIOR FILING DATE: 2000-05-22
; PRIOR APPLICATION NUMBER: US 60/218,950
; PRIOR FILING DATE: 2000-07-14
; PRIOR APPLICATION NUMBER: US 60/222,903
; PRIOR FILING DATE: 2000-08-03
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 10979
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 2271
; LENGTH: 100
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-057-475B-2271

```

```
Query Match          100.0%; Score 81; DB 15; Length 100;
Best Local Similarity 100.0%; Pred. No. 2.6e-06;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGTMYPGIADRMQK 15
   |||||
Db 46 GGTMYPGIADRMQK 60

RESULT 7
US-10-154-884B-1768
; Sequence 1768, Application US/10154884B
; Publication No. US20040005561A1
; GENERAL INFORMATION:
; APPLICANT: Gaiger, Alexander
; APPLICANT: Algate, Paul A.
; APPLICANT: Mannion, Jane
; APPLICANT: Retter, Marc W.
; APPLICANT: Corixa Corporation
; TITLE OF INVENTION: Compositions and Methods for the Detection, Diagnosis and Therapy
; TITLE OF INVENTION: Hematological Malignancies
; FILE REFERENCE: 014058-013521US
; CURRENT APPLICATION NUMBER: US/10/154,884B
; PRIOR FILING DATE: 2002-05-23
; PRIOR APPLICATION NUMBER: US 60/186,126
; PRIOR FILING DATE: 2000-03-01
; PRIOR APPLICATION NUMBER: US 60/190,479
; PRIOR FILING DATE: 2000-03-17
; PRIOR APPLICATION NUMBER: US 60/200,545
; PRIOR FILING DATE: 2000-04-27
; PRIOR APPLICATION NUMBER: US 60/200,303
; PRIOR FILING DATE: 2000-04-28
; PRIOR APPLICATION NUMBER: US 60/200,779
; PRIOR FILING DATE: 2000-04-28
; PRIOR APPLICATION NUMBER: US 60/200,999
; PRIOR FILING DATE: 2000-05-01
; PRIOR APPLICATION NUMBER: US 60/202,084
; PRIOR FILING DATE: 2000-05-04
; PRIOR APPLICATION NUMBER: US 60/206,201
; PRIOR FILING DATE: 2000-05-22
; PRIOR APPLICATION NUMBER: US 60/218,950
; PRIOR FILING DATE: 2000-07-14
; PRIOR APPLICATION NUMBER: US 60/222,903
; PRIOR FILING DATE: 2000-08-03
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 11290
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 1768
; LENGTH: 100
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-154-884B-1768

Query Match          100.0%; Score 81; DB 15; Length 100;
Best Local Similarity 100.0%; Pred. No. 2.6e-06;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGTMYPGIADRMQK 15
   |||||
Db 46 GGTMYPGIADRMQK 60

RESULT 8
US-10-154-884B-2271
; Sequence 2271, Application US/10154884B
; Publication No. US20040005561A1
; GENERAL INFORMATION:
; APPLICANT: Gaiger, Alexander
; APPLICANT: Algate, Paul A.
; APPLICANT: Mannion, Jane
; APPLICANT: Retter, Marc W.
; APPLICANT: Corixa Corporation
```

```
; TITLE OF INVENTION: Compositions and Methods for the Detection, Diagnosis and Therapy
; TITLE OF INVENTION: Hematological Malignancies
; FILE REFERENCE: 014058-013521US
; CURRENT APPLICATION NUMBER: US/10/154,884B
; PRIOR FILING DATE: 2002-05-23
; PRIOR APPLICATION NUMBER: US 60/186,126
; PRIOR FILING DATE: 2000-03-01
; PRIOR APPLICATION NUMBER: US 60/190,479
; PRIOR FILING DATE: 2000-03-17
; PRIOR APPLICATION NUMBER: US 60/200,545
; PRIOR FILING DATE: 2000-04-27
; PRIOR APPLICATION NUMBER: US 60/200,303
; PRIOR FILING DATE: 2000-04-28
; PRIOR APPLICATION NUMBER: US 60/200,779
; PRIOR FILING DATE: 2000-04-28
; PRIOR APPLICATION NUMBER: US 60/200,999
; PRIOR FILING DATE: 2000-05-01
; PRIOR APPLICATION NUMBER: US 60/202,084
; PRIOR FILING DATE: 2000-05-04
; PRIOR APPLICATION NUMBER: US 60/206,201
; PRIOR FILING DATE: 2000-05-22
; PRIOR APPLICATION NUMBER: US 60/218,950
; PRIOR FILING DATE: 2000-07-14
; PRIOR APPLICATION NUMBER: US 60/222,903
; PRIOR FILING DATE: 2000-08-03
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 11290
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 2271
; LENGTH: 100
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-154-884B-2271

Query Match          100.0%; Score 81; DB 15; Length 100;
Best Local Similarity 100.0%; Pred. No. 2.6e-06;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGTMYPGIADRMQK 15
   |||||
Db 46 GGTMYPGIADRMQK 60

RESULT 9
US-10-764-324-1768
; Sequence 1768, Application US/10764324
; Publication No. US20040175739A1
; GENERAL INFORMATION:
; APPLICANT: Gaiger, Alexander
; APPLICANT: Algate, Paul A.
; APPLICANT: Mannion, Jane
; APPLICANT: Retter, Marc
; APPLICANT: Corixa Corporation
; TITLE OF INVENTION: Compositions and Methods for the Detection, Diagnosis and Therapy
; TITLE OF INVENTION: Hematological Malignancies
; FILE REFERENCE: 014058-013520US
; CURRENT APPLICATION NUMBER: US/10/764,324
; CURRENT FILING DATE: 2004-01-23
; PRIOR APPLICATION NUMBER: US/10/040,862
; PRIOR FILING DATE: 2001-11-06
; PRIOR APPLICATION NUMBER: US 60/186,126
; PRIOR FILING DATE: 2000-03-01
; PRIOR APPLICATION NUMBER: US 60/190,479
; PRIOR FILING DATE: 2000-03-17
; PRIOR APPLICATION NUMBER: US 60/200,545
; PRIOR FILING DATE: 2000-04-27
; PRIOR APPLICATION NUMBER: US 60/200,303
; PRIOR FILING DATE: 2000-04-28
; PRIOR APPLICATION NUMBER: US 60/200,779
; PRIOR FILING DATE: 2000-04-28
; PRIOR APPLICATION NUMBER: US 60/200,999
; PRIOR FILING DATE: 2000-05-01
; PRIOR APPLICATION NUMBER: US 60/202,084
```

```
; PRIOR FILING DATE: 2000-05-04
; PRIOR APPLICATION NUMBER: US 60/206,201
; PRIOR FILING DATE: 2000-05-22
; PRIOR APPLICATION NUMBER: US 60/218,950
; PRIOR FILING DATE: 2000-07-14
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 10467
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 1768
; LENGTH: 100
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-764-324-1768

Query Match      100.0%; Score 81; DB 16; Length 100;
Best Local Similarity 100.0%; Pred. No. 2.6e-06;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGTMYPGIADRMQK 15
   |||||
Db 46 GGTMYPGIADRMQK 60

RESULT 10
US-10-764-324-2271
; Sequence 2271, Application US/10764324
; Publication No. US20040175739A1
; GENERAL INFORMATION:
; APPLICANT: Gaiger, Alexander
; APPLICANT: Algate, Paul A.
; APPLICANT: Mannion, Jane
; APPLICANT: Retter, Marc
; APPLICANT: Corixa Corporation
; TITLE OF INVENTION: Compositions and Methods for the Detection, Diagnosis and Therapy
; FILE REFERENCE: 014058-0135200S
; CURRENT APPLICATION NUMBER: US/10/764,324
; CURRENT FILING DATE: 2004-01-23
; PRIOR APPLICATION NUMBER: US/10/040,862
; PRIOR FILING DATE: 2001-11-06
; PRIOR APPLICATION NUMBER: US 60/186,126
; PRIOR FILING DATE: 2000-03-01
; PRIOR APPLICATION NUMBER: US 60/190,479
; PRIOR FILING DATE: 2000-03-17
; PRIOR APPLICATION NUMBER: US 60/200,545
; PRIOR FILING DATE: 2000-04-27
; PRIOR APPLICATION NUMBER: US 60/200,303
; PRIOR FILING DATE: 2000-04-28
; PRIOR APPLICATION NUMBER: US 60/200,779
; PRIOR FILING DATE: 2000-04-28
; PRIOR APPLICATION NUMBER: US 60/200,999
; PRIOR FILING DATE: 2000-05-01
; PRIOR APPLICATION NUMBER: US 60/202,084
; PRIOR FILING DATE: 2000-05-04
; PRIOR APPLICATION NUMBER: US 60/206,201
; PRIOR FILING DATE: 2000-05-22
; PRIOR APPLICATION NUMBER: US 60/218,950
; PRIOR FILING DATE: 2000-07-14
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 10467
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 2271
; LENGTH: 100
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-764-324-2271

Query Match      100.0%; Score 81; DB 16; Length 100;
Best Local Similarity 100.0%; Pred. No. 2.6e-06;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGTMYPGIADRMQK 15
   |||||
Db 46 GGTMYPGIADRMQK 60

US-09-423-351c-4.rapb
```

```
Db 46 GGTMYPGIADRMQK 60

RESULT 11
US-10-264-049-4060
; Sequence 4060, Application US/10264049
; Publication No. US20040005579A1
; GENERAL INFORMATION:
; APPLICANT: Birse et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
; FILE REFERENCE: P4133P1
; CURRENT APPLICATION NUMBER: US/10/264,049
; CURRENT FILING DATE: 2002-10-04
; PRIOR APPLICATION NUMBER: PCT/US01/18569
; PRIOR FILING DATE: 2001-06-07
; PRIOR APPLICATION NUMBER: US 60/209,467
; PRIOR FILING DATE: 2000-06-07
; NUMBER OF SEQ ID NOS: 4360
; SOFTWARE: PatentIn Ver. 3.1
; SEQ ID NO 4060
; LENGTH: 110
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-264-049-4060

Query Match      100.0%; Score 81; DB 15; Length 110;
Best Local Similarity 100.0%; Pred. No. 2.9e-06;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGTMYPGIADRMQK 15
   |||||
Db 36 GGTMYPGIADRMQK 50

RESULT 12
US-10-424-599-143033
; Sequence 143033, Application US/10424599
; Publication No. US20040031072A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa Thomas J
; APPLICANT: Kovalic David K
; APPLICANT: Zhou Yihua
; APPLICANT: Cao Yongwei
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(53223)B
; CURRENT APPLICATION NUMBER: US/10/424,599
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 285684
; SEQ ID NO 143033
; LENGTH: 133
; TYPE: PRT
; ORGANISM: Glycine max
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT3847_100172C.1.pep
US-10-424-599-143033

Query Match      100.0%; Score 81; DB 15; Length 133;
Best Local Similarity 100.0%; Pred. No. 3.6e-06;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGTMYPGIADRMQK 15
   |||||
Db 59 GGTMYPGIADRMQK 73

RESULT 13
US-10-264-049-4259
; Sequence 4259, Application US/10264049
; Publication No. US20040005579A1
; GENERAL INFORMATION:
; APPLICANT: Birse et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
```

```
; FILE REFERENCE: PA133PI
; CURRENT APPLICATION NUMBER: US/10/264,049
; PRIOR FILING DATE: 2002-10-04
; PRIOR APPLICATION NUMBER: PCT/US01/19569
; PRIOR FILING DATE: 2001-06-07
; PRIOR APPLICATION NUMBER: US 60/209,467
; PRIOR FILING DATE: 2000-06-07
; NUMBER OF SEQ ID NOS: 4360
; SOFTWARE: PatentIn Ver. 3.1
; SEQ ID NO 4259
; LENGTH: 145
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: MISC_FEATURE
; LOCATION: (3)
; OTHER INFORMATION: Xaa equals any of the twenty naturally occurring L-amino acids
US-10-264-049-4259
```

```
Query Match          100.0%; Score 81; DB 15; Length 145;
Best Local Similarity 100.0%; Pred. No. 3.9e-06;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY      1 GGTMYPGIADRMQK 15
        |||||
Db       71 GGTMYPGIADRMQK 85
```

```
RESULT 14
US-09-796-692-1239
; Sequence 1239, Application US/09796692
; Publication No. US20020198362A1
; GENERAL INFORMATION:
; APPLICANT: Gaiger, Alexander
; APPLICANT: Algate, Paul A.
; APPLICANT: Mannion, Jane
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE DETECTION, DIAGNOSIS AND THERAPY
; FILE REFERENCE: 2077.001200
; CURRENT APPLICATION NUMBER: US/09/796,692
; CURRENT FILING DATE: 2001-03-01
; PRIOR APPLICATION NUMBER: 60/186,126
; PRIOR FILING DATE: 2000-03-01
; PRIOR APPLICATION NUMBER: 60/190,479
; PRIOR FILING DATE: 2000-03-17
; PRIOR APPLICATION NUMBER: 60/200,545
; PRIOR FILING DATE: 2000-04-27
; PRIOR APPLICATION NUMBER: 60/200,303
; PRIOR FILING DATE: 2000-04-28
; PRIOR APPLICATION NUMBER: 60/200,779
; PRIOR FILING DATE: 2000-05-01
; PRIOR APPLICATION NUMBER: 60/202,084
; PRIOR FILING DATE: 2000-05-04
; PRIOR APPLICATION NUMBER: 60/206,201
; PRIOR FILING DATE: 2000-05-22
; PRIOR APPLICATION NUMBER: 60/218,950
; PRIOR FILING DATE: 2000-07-14
; PRIOR APPLICATION NUMBER: 60/222,903
; PRIOR FILING DATE: 2000-08-03
; PRIOR APPLICATION NUMBER: 60/223,416
; PRIOR FILING DATE: 2000-08-04
; PRIOR APPLICATION NUMBER: 60/223,378
; NUMBER OF SEQ ID NOS: 9597
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 1239
; LENGTH: 147
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-796-692-1239
```

```
Query Match          100.0%; Score 81; DB 9; Length 147;
Best Local Similarity 100.0%; Pred. No. 4e-06;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY      1 GGTMYPGIADRMQK 15
        |||||
Db       73 GGTMYPGIADRMQK 87
```

```
RESULT 15
US-09-796-692-1696
; Sequence 1696, Application US/09796692
; Publication No. US20020198362A1
; GENERAL INFORMATION:
; APPLICANT: Gaiger, Alexander
; APPLICANT: Algate, Paul A.
; APPLICANT: Mannion, Jane
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE DETECTION, DIAGNOSIS AND THERAPY
; FILE REFERENCE: 2077.001200
; CURRENT APPLICATION NUMBER: US/09/796,692
; CURRENT FILING DATE: 2001-03-01
; PRIOR APPLICATION NUMBER: 60/186,126
; PRIOR FILING DATE: 2000-03-01
; PRIOR APPLICATION NUMBER: 60/190,479
; PRIOR FILING DATE: 2000-03-17
; PRIOR APPLICATION NUMBER: 60/200,545
; PRIOR FILING DATE: 2000-04-27
; PRIOR APPLICATION NUMBER: 60/200,303
; PRIOR FILING DATE: 2000-04-28
; PRIOR APPLICATION NUMBER: 60/200,779
; PRIOR FILING DATE: 2000-04-28
; PRIOR APPLICATION NUMBER: 60/200,999
; PRIOR FILING DATE: 2000-05-01
; PRIOR APPLICATION NUMBER: 60/202,084
; PRIOR FILING DATE: 2000-05-04
; PRIOR APPLICATION NUMBER: 60/206,201
; PRIOR FILING DATE: 2000-05-22
; PRIOR APPLICATION NUMBER: 60/218,950
; PRIOR FILING DATE: 2000-07-14
; PRIOR APPLICATION NUMBER: 60/222,903
; PRIOR FILING DATE: 2000-08-03
; PRIOR APPLICATION NUMBER: 60/223,416
; PRIOR FILING DATE: 2000-08-04
; PRIOR APPLICATION NUMBER: 60/223,378
; NUMBER OF SEQ ID NOS: 9597
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 1696
; LENGTH: 147
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-796-692-1696
```

```
Query Match          100.0%; Score 81; DB 9; Length 147;
Best Local Similarity 100.0%; Pred. No. 4e-06;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY      1 GGTMYPGIADRMQK 15
        |||||
Db       73 GGTMYPGIADRMQK 87
```

```
Search completed: April 8, 2005, 12:50:58
Job time : 43.6667 secs
```



GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: April 8, 2005, 09:58:56 ; Search time 9.33333 Seconds  
(without alignments)  
154.634 Million cell updates/sec

Title: US-09-423-351C-4  
Perfect score: 81  
Sequence: 1 GGTTMPGIADRMQK 15

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 283416 seqs, 96216763 residues  
Total number of hits satisfying chosen parameters: 283416

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000  
Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : PIR 79: \*  
1: pir1.\*  
2: pir2.\*  
3: pir3.\*  
4: pir4.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	81	100.0	71	139393	alpha-actin - huma
2	81	100.0	100	214851	actin - fruit fly
3	81	100.0	125	226337	actin, skeletal mu
4	81	100.0	195	239777	actin beta - pig (
5	81	100.0	213	2461043	actin CA15 - sea s
6	81	100.0	308	2403000	actin 3 - fruit fl
7	81	100.0	328	2505430	actin beta - grass
8	81	100.0	349	225819	actin, fetal skele
9	81	100.0	362	226559	actin type 5, cyto
10	81	100.0	374	147808	actin beta - bovi
11	81	100.0	374	147808	actin gamma - bovi
12	81	100.0	374	225818	gamma-actin - huma
13	81	100.0	375	147808	actin, aortic smoo
14	81	100.0	375	147808	actin, skeletal mu
15	81	100.0	375	147808	actin beta - rat
16	81	100.0	375	147808	actin beta, cytosk
17	81	100.0	375	147808	actin beta - chick
18	81	100.0	375	147808	actin beta - huma
19	81	100.0	375	147808	actin gamma 1 - hu
20	81	100.0	375	147808	actin beta - mouse
21	81	100.0	375	147808	actin gamma - mous
22	81	100.0	375	147808	actin beta, non-mu
23	81	100.0	375	147808	actin gamma, cytos
24	81	100.0	375	147808	actin, cytosolic (
25	81	100.0	375	147808	actin 1 - Pneumocy
26	81	100.0	375	147808	actin - fission ye
27	81	100.0	375	147808	actin beta-2, cyto
28	81	100.0	375	147808	actin beta-1, cyto
29	81	100.0	375	147808	actin beta, cyto

actin beta - goose  
actin alpha, Cardi  
actin - imperfect  
actin 7 - fruit fl  
actin gamma, cytos  
actin gamma, smoot  
Actin-1A - nematod  
actin 8 - fruit fl  
actin 87E - fruit  
actin - fruit fly  
actin gamma, enter  
actin - pork tapew  
actin 4 - Caenorha  
actin 1 and actin  
actin 2 - Caenorha  
hypothetical prote

## ALIGNMENTS

### RESULT 1

139393  
alpha-actin - human (fragment)  
C:Species: Homo sapiens (man)  
C>Date: 06-Sep-1996 #sequence\_revision 06-Sep-1996 #text\_change 09-Jul-2004  
C:Accession: I39393  
R:Fuchs, E.; Kim, K.H.; Hanukoglu, I.; Tanese, N.  
Curr. Probl. Dermatol. 11, 27-44, 1983  
A:Title: The evolution and complexity of the genes encoding the cytoskeletal proteins of  
A:Reference number: I39393; MUID:84083477; PMID:6686106  
A:Accession: I39393  
A:Status: preliminary; translated from GB/EMBL/DBJ  
A:Molecule type: mRNA  
A:Residues: 1-71 <RES>  
A:Cross-references: UNIPROT:Q9UE89; GB:M28424; NID:gl78023; PIDN:AAB16906.1; PID:gl78024  
C:Superfamily: actin

Query Match 100.0%; Score 81; DB 2; Length 71;  
Best Local Similarity 100.0%; Pred. No. 7.8e-07; Mismatches 0; Indels 0; Gaps 0;  
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GGTTMPGIADRMQK 15  
Db 40 GGTTMPGIADRMQK 54

### RESULT 2

214851  
actin - fruit fly (Drosophila melanogaster)  
C:Species: Drosophila melanogaster  
C>Date: 20-Feb-1995 #sequence\_revision 20-Feb-1995 #text\_change 09-Jul-2004  
C:Accession: S14851  
R:Burn, T.C.; Tobin, S.L.  
submitted to the EMBL Data Library, October 1990  
A:Reference number: S14851  
A:Accession: S14851  
A:Molecule type: DNA  
A:Residues: 1-100 <BUR>  
A:Cross-references: UNIPROT:P02572; EMBL:X54848; NID:g7549; PIDN:CAA38618.1; PID:g7550  
C:Genetics:

A:Gene: FlyBase:Act42A  
A:Cross-references: FlyBase:FBgn0000043  
C:Superfamily: actin

C:Keywords: cell motility; cytoskeleton; microfilament; mitosis; structural protein

Query Match 100.0%; Score 81; DB 2; Length 100;  
Best Local Similarity 100.0%; Pred. No. 1.1e-06; Mismatches 0; Indels 0; Gaps 0;  
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GGTTMPGIADRMQK 15  
Db 26 GGTTMPGIADRMQK 40

RESULT 3  
 A26337  
 actin, skeletal muscle - Iberian ribbed newt (fragment)  
 C:Species: Pleurodeles waltlii (Iberian ribbed newt)  
 C>Date: 19-Nov-1988 #sequence\_revision 19-Nov-1988 #text\_change 09-Jul-2004  
 C:Accession: A26337  
 R:Khrstchatisky, M.; Pontes, M.  
 J. Mol. Biol. 193, 409-412, 1987  
 A:Title: There is an alpha-actin skeletal muscle-specific gene in a salamander (Pleurodeles waltlii)  
 A:Reference number: A26337; MUID:87254218; PMID:3453122  
 A:Accession: A26337  
 A:Molecule type: DNA  
 A:Residues: 1-125 <KHR>  
 A:CROSS-references: UNIPROT:P10994; GB:X05106; NID:964248; PID:CAA28753.1; PID:964249  
 C:Superfamily: actin  
 C:Keywords: muscle; skeletal muscle  
  
 Query Match 100.0%; Score 81; DB 2; Length 125;  
 Best Local Similarity 100.0%; Pred. No. 1.4e-06; Mismatches 0; Indels 0; Gaps 0;  
 Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
  
 QY 1 GGTMYPGIADRMQK 15  
 |||||  
 DB 51 GGTMYPGIADRMQK 65  
  
 RESULT 4  
 S39777  
 actin beta - pig (fragments)  
 C:Species: Sus scrofa domestica (domestic pig)  
 C>Date: 07-Oct-1994 #sequence\_revision 12-Apr-1996 #text\_change 09-Jul-2004  
 C:Accession: S39777  
 R:Adamski, J.; Hussen, B.; Thole, H.H.; Groeschel-Stewart, U.; Jungblut, P.W.  
 Biochem. J. 296, 797-802, 1993  
 A:Title: Linkage of 17-beta-estradiol dehydrogenase to actin by epsilon-(gamma-glutamyl-1-ATP)  
 A:Reference number: S39777; MUID:94107247; PMID:8280079  
 A:Accession: S39777  
 A:Molecule type: protein  
 A:Residues: 1-20;21-52;53-79;80-85;86-95;96-116;117-151;152-159;160-195 <ADA>  
 A:CROSS-references: UNIPROT:Q7M3B0  
 C:Superfamily: actin  
  
 Query Match 100.0%; Score 81; DB 2; Length 195;  
 Best Local Similarity 100.0%; Pred. No. 2.2e-06; Mismatches 0; Indels 0; Gaps 0;  
 Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
  
 QY 1 GGTMYPGIADRMQK 15  
 |||||  
 DB 126 GGTMYPGIADRMQK 140  
  
 RESULT 5  
 A61043  
 actin CA15 - sea squirt (Styela clava) (fragments)  
 C:Species: Styela clava  
 C>Date: 31-Dec-1993 #sequence\_revision 31-Dec-1993 #text\_change 09-Jul-2004  
 C:Accession: A61043  
 R:Beach, R.L.; Jeffery, W.R.  
 Dev. Genet. 11, 2-14, 1990  
 A:Title: Temporal and spatial expression of a cytoskeletal actin gene in the ascidian Styela clava  
 A:Reference number: A61043; MUID:90298580; PMID:2361333  
 A:Accession: A61043  
 A:Status: not compared with conceptual translation  
 A:Molecule type: mRNA  
 A:Residues: 1-213 <BEA>  
 A:CROSS-references: UNIPROT:Q7M3Y7  
 C:Comment: This sequence is expressed in cells undergoing rapid cell division.  
 C:Superfamily: actin  
 C:Keywords: cytoskeleton; methylated amino acid; mitosis; structural protein  
 F,7/3/Modified site: 3'-methylhistidine (His) #status predicted

Query Match 100.0%; Score 81; DB 2; Length 213;  
 Best Local Similarity 100.0%; Pred. No. 2.4e-06; Mismatches 0; Indels 0; Gaps 0;  
 Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
  
 QY 1 GGTMYPGIADRMQK 15  
 |||||  
 DB 139 GGTMYPGIADRMQK 153  
  
 RESULT 6  
 A03000  
 actin 3 - fruit fly (Drosophila melanogaster) (fragments)  
 C:Species: Drosophila melanogaster  
 C>Date: 31-Mar-1991 #sequence\_revision 31-Mar-1991 #text\_change 09-Jul-2004  
 C:Accession: A03000  
 R:Fyberg, E.A.; Bond, B.J.; Hershey, N.D.; Mixter, K.S.; Davidson, N.  
 Cell 24, 107-116, 1981  
 A:Title: The actin genes of Drosophila: protein coding regions are highly conserved but  
 A:Reference number: A03000; MUID:81210174; PMID:6263481  
 A:Accession: A03000  
 A:Molecule type: DNA  
 A:Residues: 1-308 <FYR>  
 A:CROSS-references: UNIPROT:P02572  
 A:Note: there are 68 unsequenced residues between positions 160 and 161. Partial sequence  
 A:Note: the authors translated the codon GTT for residue 263 as Ile  
 C:Genetics:  
 A:Gene: FlyBase:Act42A  
 A:CROSS-references: FlyBase:FBgn0000043  
 A:Map position: 42A  
 C:Superfamily: actin  
 C:Keywords: methylated amino acid  
 F,7/4/Modified site: 3'-methylhistidine (His) #status predicted

Query Match 100.0%; Score 81; DB 2; Length 308;  
 Best Local Similarity 100.0%; Pred. No. 3.6e-06; Mismatches 0; Indels 0; Gaps 0;  
 Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
  
 QY 1 GGTMYPGIADRMQK 15  
 |||||  
 DB 234 GGTMYPGIADRMQK 248  
  
 RESULT 7  
 S05430  
 actin beta - grass carp  
 C:Species: Ctenopharyngodon idella (grass carp)  
 C>Date: 07-Jun-1990 #sequence\_revision 07-Jun-1990 #text\_change 09-Jul-2004  
 C:Accession: S05430  
 R:Liu, Z.; Zhu, Z.; Roberg, K.; Faras, A.J.; Guise, K.S.; Kapuscinski, A.R.; Hackett, P.  
 Nucleic Acids Res. 17, 5850, 1989  
 A:Title: The beta-actin gene of carp (Ctenopharyngodon idella).  
 A:Reference number: S05430; MUID:89345185; PMID:2762162  
 A:Accession: S05430  
 A:Status: translation not shown  
 A:Molecule type: DNA  
 A:Residues: 1-328 <LIU>  
 A:CROSS-references: UNIPROT:P83751; EMBL:M25013  
 C:Genetics:  
 A:Introns: 41/3; 121/3; 268/1  
 C:Superfamily: actin  
 C:Keywords: cytoskeleton; methylated amino acid  
 F,7/3/Modified site: 3'-methylhistidine (His) #status predicted

Query Match 100.0%; Score 81; DB 2; Length 328;  
 Best Local Similarity 100.0%; Pred. No. 3.8e-06; Mismatches 0; Indels 0; Gaps 0;  
 Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
  
 QY 1 GGTMYPGIADRMQK 15  
 |||||  
 DB 301 GGTMYPGIADRMQK 315  
  
 RESULT 8

B25819  
actin, fetal skeletal/adult cardiac muscle - mouse (fragment)  
C:Species: Mus musculus (house mouse)  
C:Date: 23-Aug-1987 #sequence\_revision 23-Aug-1987 #text\_change 09-Jul-2004  
C:Accession: B25819  
R:Alonso, S.; Minty, A.; Bourlet, Y.; Buckingham, M.  
J. Mol. Evol. 23, 11-22, 1986  
A:Title: Comparison of three actin-coding sequences in the mouse; evolutionary relations  
A:Reference number: A25819; MUID:86200234; PMID:3084797  
A:Accession: B25819  
A:Molecule type: mRNA  
A:Residues: 1-349 <ALO>  
A:Cross-references: UNIPROT:Q61275; GB:J00381; GB:M10652; NID:G49869; PIDN:CA  
C:Superfamily: actin  
C:Keywords: cardiac muscle; heart; methylated amino acid; muscle; skeletal muscle  
F:47/Modified site: 3'-methylhistidine (His) #status predicted

Query Match 100.0%; Score 81; DB 2; Length 349;  
Best Local Similarity 100.0%; Pred. No. 4.1e-06;  
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GGTMYPGIADRMQK 15  
Db 275 GGTMYPGIADRMQK 289  
|||||

RESULT 9  
A26559  
actin type 5, cytosolic - chicken  
C:Species: Gallus gallus (chicken)  
C:Date: 09-Sep-1987 #sequence\_revision 09-Sep-1987 #text\_change 05-Dec-1997  
C:Accession: A26559  
R:Bergsma, D.J.; Chang, K.S.; Schwartz, R.J.  
Mol. Cell. Biol. 5, 1151-1162, 1985  
A:Reference number: A26559; MUID:85213487; PMID:4000121  
A:Accession: A26559  
A:Molecule type: DNA  
A:Residues: 1-362 <BER>  
C:Superfamily: actin  
C:Keywords: cytosol; methylated amino acid  
F:74/Modified site: 3'-methylhistidine (His) #status predicted

Query Match 100.0%; Score 81; DB 2; Length 362;  
Best Local Similarity 100.0%; Pred. No. 4.2e-06;  
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GGTMYPGIADRMQK 15  
Db 302 GGTMYPGIADRMQK 316  
|||||

RESULT 10  
ATBOG  
actin beta - bovine (tentative sequence)  
C:Species: Bos primigenius taurus (cattle)  
C:Date: 30-Sep-1991 #sequence\_revision 30-Sep-1991 #text\_change 09-Jul-2004  
C:Accession: E14185; A39105; A02999; A14185  
R:Vandekerckhove, J.; Weber, K.  
Eur. J. Biochem. 90, 451-462, 1978  
A:Title: Actin amino-acid sequences. Comparison of actins from calf thymus, bovine brain  
A:Reference number: A14185; MUID:79045349; PMID:213279  
A:Accession: E14185  
A:Molecule type: protein  
A:Residues: 1-374 <VAN>  
A:Cross-references: UNIPROT:P60712  
A:Note: only peptides that differed in composition from the corresponding peptides of ra  
R:Degen, J.L.; Neubauer, M.G.; Degen, S.J.F.; Seyfried, C.B.; Morris, D.R.  
J. Biol. Chem. 258, 12153-12162, 1983  
A:Title: Regulation of protein synthesis in mitogen-activated bovine lymphocytes. Analysis  
A:Reference number: A39105; MUID:84032385; PMID:6195151  
A:Accession: A39105  
A:Molecule type: mRNA  
A:Residues: 76-227;344-374 <DEG>

A:Cross-references: GB:K00622; GB:K00623  
A:Note: actins beta and gamma were not distinguished in this study  
C:Comment: Vertebrate nonmuscle cells contain two highly conserved cytoskeletal actins,  
C:Superfamily: actin  
C:Keywords: blocked amino end; cell motility; cytoskeleton; methylated amino acid; micro  
F:1/Modified site: blocked amino end (Asp) (probably acetylated) #status experimental  
F:72/Modified site: 3'-methylhistidine (His) #status predicted

Query Match 100.0%; Score 81; DB 1; Length 374;  
Best Local Similarity 100.0%; Pred. No. 4.4e-06;  
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GGTMYPGIADRMQK 15  
Db 300 GGTMYPGIADRMQK 314  
|||||

RESULT 11  
ATBOG  
actin gamma - bovine (tentative sequence)  
C:Species: Bos primigenius taurus (cattle)  
C:Date: 30-Sep-1991 #sequence\_revision 30-Sep-1991 #text\_change 09-Jul-2004  
C:Accession: E14185; A02999  
R:Vandekerckhove, J.; Weber, K.  
Eur. J. Biochem. 90, 451-462, 1978  
A:Title: Actin amino-acid sequences. Comparison of actins from calf thymus, bovine brain  
A:Reference number: A14185; MUID:79045349; PMID:213279  
A:Accession: E14185  
A:Molecule type: protein  
A:Residues: 1-374 <VAN>  
A:Cross-references: UNIPROT:P02571  
A:Note: only peptides that differed in composition from the corresponding peptides of ra  
C:Comment: Vertebrate nonmuscle cells contain two highly conserved cytoskeletal actins,  
C:Superfamily: actin  
C:Keywords: blocked amino end; cell motility; cytoskeleton; methylated amino acid; micro  
F:1/Modified site: blocked amino end (Glu) (probably acetylated) #status experimental  
F:72/Modified site: 3'-methylhistidine (His) #status predicted

Query Match 100.0%; Score 81; DB 1; Length 374;  
Best Local Similarity 100.0%; Pred. No. 4.4e-06;  
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GGTMYPGIADRMQK 15  
Db 300 GGTMYPGIADRMQK 314  
|||||

RESULT 12  
JC5818  
gamma-actin - human  
C:Species: Homo sapiens (man)  
C:Date: 20-Feb-1998 #sequence\_revision 13-Mar-1998 #text\_change 09-Jul-2004  
C:Accession: JC5818; PC4501  
R:Hauschildt, S.; Schwarz, C.; Heine, H.; Ulmer, A.J.; Flad, H.D.; Rietschel, E.T.; Jens  
Biochem. Biophys. Res. Commun. 241, 670-674, 1997  
A:Title: Actin: A target of lipopolysaccharide-induced phosphorylation in human monocytes  
A:Reference number: JC5818; MUID:98096379; PMID:9434766  
A:Accession: JC5818  
A:Molecule type: protein  
A:Residues: 1-374 <HAU>  
A:Cross-references: UNIPROT:P02571  
A:Experimental source: monocyte  
A:Accession: PC4501  
A:Molecule type: protein  
A:Residues: 1-61;84-112;147-190;196-209;215-253;335-358 <HA2>  
A:Experimental source: monocyte  
C:Comment: This protein is involved in a signal transduction that eventually leads to mo  
C:Superfamily: actin

Query Match 100.0%; Score 81; DB 2; Length 374;  
Best Local Similarity 100.0%; Pred. No. 4.4e-06;  
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGTMYPGIADRMQK 15  
 |||||  
 Db 300 GGTMYPGIADRMQK 314

RESULT 13  
 A:Title: actin, aortic smooth muscle - bovine  
 C:Species: Bos primigenius taurus (cattle)  
 C>Date: 30-Jun-1987 #sequence\_revision 30-Jun-1987 #text\_change 06-Sep-1996  
 C:Accession: A02997; S13480  
 R:Vandekerckhove, J.; Weber, K.  
 A:Title: The complete amino acid sequence of actins from bovine aorta, bovine heart, bovine thymus.  
 A:Reference number: A02997; PMID:80047657; PMID:499690  
 A:Accession: A02997  
 A:Molecule type: protein  
 A:Residues: 1-375 <VAN>  
 R:Zevgoulis, V.G.; Sotiropoulos, T.G.; Evangelopoulos, A.E.  
 Biochim. Biophys. Acta 1091, 222-230, 1991  
 A:Title: Phosphorylase kinase from bovine stomach smooth muscle: a Ca (2+) -dependent protein  
 A:Reference number: S13480; PMID:91137633; PMID:1995080  
 A:Accession: S13480  
 A:Molecule type: protein  
 A:Residues: 40-49 <ZEV>  
 A:Experimental source: stomach  
 A:Note: This material appears to be actin of aortic smooth muscle type or a related molecule  
 C:Superfamily: actin  
 C:Keywords: acetylated amino end; methylated amino acid; muscle contraction  
 F:1/Modified site: acetylated amino end (Glu) #status predicted  
 F:73/Modified site: 3'-methylhistidine (His) #status experimental

Query Match 100.0%; Score 81; DB 1; Length 375;  
 Best Local Similarity 100.0%; Pred. No. 4.4e-06;  
 Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGTMYPGIADRMQK 15  
 |||||  
 Db 301 GGTMYPGIADRMQK 315

RESULT 14  
 A:Title: actin, skeletal muscle - rabbit  
 N:Alternate names: F-actin  
 C:Species: Oryctolagus cuniculus (domestic rabbit)  
 C>Date: 30-Sep-1991 #sequence\_revision 30-Sep-1991 #text\_change 07-May-1999  
 C:Accession: A92182; A14185; S65873; S70610; S44393; A02994  
 R:Collins, J.H.; Elzinga, M.  
 J. Biol. Chem. 250, 5915-5920, 1975  
 A:Title: The primary structure of actin from rabbit skeletal muscle. Completion and analysis  
 A:Reference number: A92182; PMID:75211334; PMID:1150665  
 A:Accession: A92182  
 A:Molecule type: protein  
 A:Residues: 1-2, 'T', 'D', '6-11', 'D', '13-73', 'W', '74-78, 80-234, 236-308, 'T', 310-375 <COL>  
 A:Note: this is the final paper in a series  
 A:Note: this sequence has been revised in references A14185 and A90406  
 R:Vandekerckhove, J.; Weber, K.  
 Eur. J. Biochem. 90, 451-462, 1978  
 A:Title: Actin amino-acid sequences. Comparison of actins from calf thymus, bovine brain  
 A:Reference number: A14185; PMID:79045349; PMID:213279  
 A:Accession: A14185  
 A:Molecule type: protein  
 A:Residues: 1-18; 69-84 <VAN>  
 R:Li, R.C.; Elzinga, M.  
 Biochemistry 16, 5801-5806, 1977  
 A:Title: Partial amino acid sequence of brain actin and its homology with muscle actin.  
 A:Reference number: A90406; PMID:78060866; PMID:588555  
 A:Contents: annotation  
 A:Note: residue 235 has been added and residue 309 has been revised  
 R:Vandat, A.; Miller, C.; Phillips, M.; Muhlrad, A.; Reisler, E.  
 FEBS Lett. 365, 149-151, 1995

A:Title: A novel 27/16 kDa form of subtilisin cleaved actin: structural and functional characterization  
 A:Reference number: S65873; PMID:95300963; PMID:7781768  
 A:Accession: S65873  
 A:Molecule type: protein  
 A:Residues: 235-241 <VAH>  
 R:Strzelecka-Golaszewska, H.; Wozniak, A.; Hult, T.; Lindberg, U.  
 Biochem. J. 316, 713-721, 1996  
 A:Title: Effects of the type of divalent cation, Ca2+ or Mg2+, bound at the high-affinity binding site on the rate of actin polymerization  
 A:Reference number: S70610; PMID:96265033; PMID:8670143  
 A:Accession: S70610  
 A:Molecule type: protein  
 A:Residues: 48-54; 68-72; 235-243 <STR>  
 A:Experimental source: skeletal muscle  
 R:Bertrand, R.; Derancourt, J.; Kassab, R.  
 FEBS Lett. 345, 113-119, 1994  
 A:Title: The covalent maleimidebenzoyl-actin-myosin head complex. Cross-linking of the 50 kDa myosin head with a maleimide derivative  
 A:Reference number: S44393; PMID:94259162; PMID:8200441  
 A:Accession: S44393  
 A:Molecule type: protein  
 A:Residues: 48-64 <BER>  
 A:Experimental source: skeletal muscle  
 C:Superfamily: actin  
 C:Keywords: acetylated amino end; ATP binding; methylated amino acid; muscle contraction  
 F:1-375/Product: actin #status experimental <MAT>  
 F:1/Modified site: acetylated amino end (Asp) #status experimental  
 F:73/Modified site: 3'-methylhistidine (His) #status experimental

Query Match 100.0%; Score 81; DB 1; Length 375;  
 Best Local Similarity 100.0%; Pred. No. 4.4e-06;  
 Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGTMYPGIADRMQK 15  
 |||||  
 Db 301 GGTMYPGIADRMQK 315

RESULT 15  
 A:Title: actin beta - rat  
 C:Species: Rattus norvegicus (Norway rat)  
 C>Date: 13-Jun-1983 #sequence\_revision 13-Jun-1983 #text\_change 22-Jun-1999  
 C:Accession: A38571; A02999  
 R:Nudel, U.; Zakut, R.; Shani, M.; Neuman, S.; Levy, Z.; Yaffe, D.  
 Nucleic Acids Res. 11, 1759-1771, 1983  
 A:Title: The nucleotide sequence of the rat cytoplasmic beta-actin gene.  
 A:Reference number: A38571; PMID:83168920; PMID:6300777  
 A:Accession: A38571  
 A:Molecule type: DNA  
 A:Residues: 1-375 <NUD>  
 A:Cross-references: GB:J00691; NID:G202653; PIDN:AAA40657.1; PID:G202654  
 C:Comment: Vertebrate nonmuscle cells contain two highly conserved cytoskeletal actins, actin beta and actin gamma.  
 C:Genetics: 41/3; 121/3; 268/1; 328/3  
 C:Introns: 41/3; 121/3; 268/1; 328/3  
 C:Superfamily: actin  
 C:Keywords: cell motility; cytoskeleton; methylated amino acid; microfilament; mitosis; F-2-375/Product: actin beta #status predicted <MAT>  
 F:73/Modified site: 3'-methylhistidine (His) #status predicted

Query Match 100.0%; Score 81; DB 1; Length 375;  
 Best Local Similarity 100.0%; Pred. No. 4.4e-06;  
 Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGTMYPGIADRMQK 15  
 |||||  
 Db 301 GGTMYPGIADRMQK 315

Search completed: April 8, 2005, 10:53:05  
 Job time : 9.33333 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

# OM protein - protein search, using sw model

Run on: April 8, 2005, 09:58:56 ; Search time 57.8667 Seconds  
(without alignments)  
132.739 Million cell updates/sec

Title: US-09-423-351C-4

Perfect score: 81

Sequence: 1 GGTTMYPGIADRMQK 15

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1612378 seqs, 512079187 residues

Total number of hits satisfying chosen parameters: 1612378

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : UniProt\_03:\*

1: uniprot\_sprot:\*

2: uniprot\_trembl:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	81	100.0	60	Q95164	Q95164 canis famil
2	81	100.0	71	Q9UE89	Q9UE89 homo sapien
3	81	100.0	93	Q7ZZ20	Q7ZZ20 anas platyr
4	81	100.0	96	Q9UX51	Q9UX51 fundulus he
5	81	100.0	100	Q9TYD9	Q9TYD9 styela clav
6	81	100.0	103	Q95YG5	Q95YG5 apis cerana
7	81	100.0	106	Q80UA7	Q80UA7 marmota mon
8	81	100.0	107	Q8J1Q2	Q8J1Q2 anguilla ja
9	81	100.0	109	Q91877	Q91877 diprion pin
10	81	100.0	109	Q91881	Q91881 diadromus p
11	81	100.0	110	Q76493	Q76493 bemisia arg
12	81	100.0	111	Q6R127	Q6R127 boophilus m
13	81	100.0	111	Q6R128	Q6R128 boophilus m
14	81	100.0	113	Q6Y260	Q6Y260 pagrus majo
15	81	100.0	118	Q9DF22	Q9DF22 scyllorhinu
16	81	100.0	123	Q6UIS1	Q6UIS1 macaca mula
17	81	100.0	123	Q6UIS2	Q6UIS2 pan troglod
18	81	100.0	124	Q6B9W6	Q6B9W6 parastrongy
19	81	100.0	124	Q95319	Q95319 sus scrofa
20	81	100.0	124	Q9DFK3	Q9DFK3 gillichthys
21	81	100.0	125	1 ACTS PLEWA	10994 pleurodeles
22	81	100.0	129	Q9NAY5	Q9NAY5 papilio can
23	81	100.0	131	Q9XSB1	Q9XSB1 tupiaa glis
24	81	100.0	133	Q96104	Q96104 apis mellif
25	81	100.0	134	Q99NC5	Q99NC5 mus musculu
26	81	100.0	134	Q99NC6	Q99NC6 rattus norv
27	81	100.0	141	Q9NAY6	Q9NAY6 papilio gla
28	81	100.0	141	Q6TFS8	Q6TFS8 squalus aca
29	81	100.0	141	Q6Q6W9	Q6Q6W9 ictalurus p
30	81	100.0	145	Q6B9V8	Q6B9V8 strongyloid
31	81	100.0	149	Q6PTC1	Q6PTC1 sitobion av

32 81 100.0 149 2 Q6PTC2 Q6PTC2 sitobion av  
33 81 100.0 159 2 Q96DE1 Q96DE1 homo sapien  
34 81 100.0 159 2 Q9DG95 Q9DG95 pagrus majo  
35 81 100.0 165 2 Q96FUG Q96FUG homo sapien  
36 81 100.0 170 2 Q96ER4 Q96ER4 chironomus  
37 81 100.0 171 2 Q6GX56 Q6GX56 lepisosteus  
38 81 100.0 171 2 Q7ZW59 Q7ZW59 rana cateeb  
39 81 100.0 176 2 Q96CV5 Q96CV5 drosophila  
40 81 100.0 176 2 Q96FV8 Q96FV8 drosophila  
41 81 100.0 181 2 Q6XJ33 Q6XJ33 drosophila  
42 81 100.0 183 2 Q9YGY6 Q9YGY6 orochromis  
43 81 100.0 195 2 Q7M3B0 Q7M3B0 sus scrofa  
44 81 100.0 207 2 Q6SQL9 Q6SQL9 helicon pel  
45 81 100.0 213 2 Q7M3Y7 Q7M3Y7 styela clav

## ALIGNMENTS

### RESULT 1

Q95164 PRELIMINARY; PRT; 60 AA.  
AC Q95164;  
DT 01-FEB-1997 (Tremblrel. 02, Created)  
DT 01-FEB-1997 (Tremblrel. 02, Last sequence update)  
DT 01-JUN-2003 (Tremblrel. 24, Last annotation update)  
DE Beta-actin (Fragment).  
OS Canis familiaris (Dog).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Carnivora; Fissipedia; Canidae; Canis.  
OX NCBI\_TaxID=9615;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=97008128; PubMed=8855304; DOI=10.1073/pnas.93.20.11035;  
RA Lee Y.A., Liang C.S., Lee M.A., Lindpaintner K.;  
RT "Local stress, not systemic factors, regulate gene expression of the  
cardiac renin-angiotensin system in vivo: a comprehensive study of all  
its components in the dog."  
RL Proc. Natl. Acad. Sci. U.S.A. 93:11035-11040(1996).  
DR EMBL; U67202; AAC48640.1; -  
DR HSSP; P02568; IMA9.  
DR GO; GO:0015629; C:actin cytoskeleton; IEA.  
DR GO; GO:0005200; F:structural constituent of cytoskeleton; IEA.  
DR InterPro; IPR004000; Actin\_like.  
DR Pfam; PF00022; Actin; 1.  
FT NON\_TER 1  
FT NON\_TER 60  
SQ SEQUENCE 60 AA; 6744 MW; AA02DE065D32D7C4 CRC64;  
Query Match 100.0%; Score 81; DB 2; Length 60;  
Best Local Similarity 100.0%; Pred. No. 5e-06;  
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

### RESULT 2

Q9UE89 PRELIMINARY; PRT; 71 AA.  
AC Q9UE89;  
DT 01-MAY-2000 (Tremblrel. 13, Created)  
DT 01-MAY-2000 (Tremblrel. 13, Last sequence update)  
DT 01-JUN-2003 (Tremblrel. 24, Last annotation update)  
DE Beta-actin (Fragment).  
OS Homo sapiens (Human).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
OX NCBI\_TaxID=9606;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC TISSUE=ForeSkin;  
Qy 1 GGTTMYPGIADRMQK 15  
Db 20 GGTTMYPGIADRMQK 34

```

RX MEDLINE=84083477; PubMed=6686106;
RA Fuchs E., Kim K.H., Hanukoglu I., Tanase N.;
RT "The evolution and complexity of the genes encoding the cytoskeletal
RL proteins of human epidermal cells.";
RL Curr. Probl. Dermatol. 11:27-44(1983).
DR EMBL; M28424; AAB16906.1; -.
DR PIR; I39393; I39393.
DR HSSP; P02568; 1MA9.
DR GO; GO:0015629; C:actin cytoskeleton; IEA.
DR GO; GO:0005200; F:structural constituent of cytoskeleton; IEA.
DR InterPro; IPR004000; Actin_like.
DR Pfam; PF00022; Actin; 1.
FT NON TER 1
FT NON TER 71
SQ SEQUENCE 71 AA; 7799 MW; D3886403C855CF70 CRC64;

Query Match 100.0%; Score 81; DB 2; Length 71;
Best Local Similarity 100.0%; Pred. No. 5.9e-06;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGTMYPGIADRMQK 15
Db 40 GGTMYPGIADRMQK 54

RESULT 3
Q7ZZZ0
ID Q7ZZZ0 PRELIMINARY; PRT; 93 AA.
AC Q7ZZZ0;
DT 01-JUN-2003 (TrEMBLrel. 24, Created)
DT 01-JUN-2003 (TrEMBLrel. 24, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Beta-actin (Fragment).
OS Anas platyrhynchos (Domestic duck).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Archosauria; Aves; Neognathae; Anseriformes; Anatidae; Anas.
OX NCBI_TaxID=8839;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Spleen;
RA Sreekumar E., Premraj A., Rasool T.J.;
RL Submitted (MAR-2003) to the EMBL/GenBank/DBJ databases.
DR EMBL; AY251275; AAP04479.1; -.
DR HSSP; P02568; 1QZ5.
DR GO; GO:0005884; C:actin filament; IEA.
DR GO; GO:0003774; F:motor activity; IEA.
DR GO; GO:0005200; F:structural constituent of cytoskeleton; IEA.
DR InterPro; IPR004001; Actin.
DR InterPro; IPR004000; Actin_like.
DR Pfam; PF00022; Actin; 1.
DR PROSITE; PS00432; ACTINS_2; 1.
FT NON TER 1
FT NON TER 93
SQ SEQUENCE 93 AA; 10444 MW; F867ADF144D80523 CRC64;

Query Match 100.0%; Score 81; DB 2; Length 93;
Best Local Similarity 100.0%; Pred. No. 7.8e-06;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGTMYPGIADRMQK 15
Db 19 GGTMYPGIADRMQK 33

RESULT 4
Q90X51
ID Q90X51 PRELIMINARY; PRT; 96 AA.
AC Q90X51;
DT 01-DEC-2001 (TrEMBLrel. 19, Created)
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE Beta-actin (Fragment).
OS Fundulus heteroclitus (Killifish) (Mummichog).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

```

```

OC Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;
OC Acanthomorpha; Acanthopterygii; Percomorpha; Atherinomorpha;
OC Cyprinodontiformes; Fundulidae; Fundulus.
OX NCBI_TaxID=8078;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Liver;
RA Roling J.A., Baldwin W.S.;
RL Submitted (OCT-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF355092; AAL29465.1; -.
DR HSSP; P02568; 1QZ5.
DR GO; GO:0005884; C:actin filament; IEA.
DR GO; GO:0003774; F:motor activity; IEA.
DR GO; GO:0005200; F:structural constituent of cytoskeleton; IEA.
DR Pfam; PF00022; Actin; 1.
DR PROSITE; PS00432; ACTINS_2; 1.
FT NON TER 1
FT NON TER 96
SQ SEQUENCE 96 AA; 10798 MW; 7440DA6AE5A070C4 CRC64;

Query Match 100.0%; Score 81; DB 2; Length 96;
Best Local Similarity 100.0%; Pred. No. 8.1e-06;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGTMYPGIADRMQK 15
Db 33 GGTMYPGIADRMQK 47

RESULT 5
Q9TVD9
ID Q9TVD9 PRELIMINARY; PRT; 100 AA.
AC Q9TVD9;
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE Alpha-muscle actin (Fragment).
GN Name=Tb24;
OS Styela clava (Sea squirt).
OC Eukaryota; Metazoa; Chordata; Urochordata; Ascidiacea;
OC Stolidobranchia; Styelidae; Styela.
OX NCBI_TaxID=7725;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=92249656; PubMed=1577198;
RA Beach R.L., Jeffery W.R.;
RT "Multiple actin genes encoding the same alpha-muscle isoform are
RL expressed during ascidian development.";
RL Dev. Biol. 151:55-66(1992).
DR EMBL; L13788; AAA29846.1; -.
DR HSSP; P02568; 1QZ5.
DR GO; GO:0005884; C:actin filament; IEA.
DR GO; GO:0003774; F:motor activity; IEA.
DR GO; GO:0005200; F:structural constituent of cytoskeleton; IEA.
DR InterPro; IPR004001; Actin.
DR InterPro; IPR004000; Actin_like.
DR Pfam; PF00022; Actin; 1.
DR PROSITE; PS00432; ACTINS_2; 1.
FT NON TER 1
FT NON TER 100
SQ SEQUENCE 100 AA; 11264 MW; 108CE34548BAB5E CRC64;

Query Match 100.0%; Score 81; DB 2; Length 100;
Best Local Similarity 100.0%; Pred. No. 8.4e-06;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGTMYPGIADRMQK 15
Db 26 GGTMYPGIADRMQK 40

RESULT 6
Q95YG5
ID Q95YG5 PRELIMINARY; PRT; 103 AA.

```

```
AC Q95YG5;
DT 01-DEC-2001 (TrEMBLrel. 19, Created)
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE Actin (Fragment).
OS Apis cerana (Indian honeybee).
OC Neoptera; Hymenoptera; Hexapoda; Insecta; Pterygota;
OC Eukaryota; Metazoa; Arthropoda;
OC Apidae; Apis.
OX NCBI_TaxID=7461;
RN [1]
RP SEQUENCE FROM N.A.
RA Shimizu I., Kawai Y., Taniguchi M., Aoki S.;
RT "Circadian rhythm and cDNA cloning of the clock gene period in the
RT honeybee Apis cerana japonica.";
RL Zool. Sci. 18:779-789 (2001).
DR EMBL; AB072495; BAB69483.1; -.
DR HSSP; P10983; 1D4X.
DR GO; GO:0015629; C:actin cytoskeleton; IEA.
DR GO; GO:0005200; F:structural constituent of cytoskeleton; IEA.
DR InterPro; IPR004000; Actin_like.
DR Pfam; PF00022; Actin; 1.
FT NON_TER 1
FT NON_TER 103
SQ SEQUENCE 103 AA; 11520 MW; 5203DA3630FCD807 CRC64;

Query Match 100.0%; Score 81; DB 2; Length 103;
Best Local Similarity 100.0%; Pred. No. 8.7e-06;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GGTTMPGFIADRMQK 15
Db 66 GGTTMPGFIADRMQK 80

RESULT 7
Q80UA7 PRELIMINARY; PRT; 106 AA.
AC Q80UA7;
DT 01-JUN-2003 (TrEMBLrel. 24, Created)
DT 01-JUN-2003 (TrEMBLrel. 24, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Beta-actin (Fragment).
OS Marmota monax (Woodchuck).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Sciuridae; Sciurinae;
OC Marmota.
OX NCBI_TaxID=9995;
RN [1]
RP SEQUENCE FROM N.A.
RA MEDLINE=22554704; PubMed=12668969; DOI=10.1053/jhep.2003.50154;
RA Wang Y., Menne S., Jacob J.R., Tennant B.C., Gerin J.L., Cole P.J.;
RT "Role of type 1 versus type 2 immune responses in liver during the
RT onset of chronic woodchuck hepatitis virus infection.";
RL Hepatology 37:771-780 (2003).
DR EMBL; AY170121; AAO39434.1; -.
DR HSSP; P10983; 1D4X.
DR GO; GO:0005884; C:actin filament; IEA.
DR GO; GO:0003774; F:motor activity; IEA.
DR GO; GO:0005200; F:structural constituent of cytoskeleton; IEA.
DR InterPro; IPR004001; Actin.
DR InterPro; IPR004000; Actin_like.
DR Pfam; PF00022; Actin; 1.
DR PROSITE; PS00432; ACTINS_2; 1.
FT NON_TER 1
FT NON_TER 1
SQ SEQUENCE 106 AA; 11864 MW; 9PFD3145651FE343 CRC64;

Query Match 100.0%; Score 81; DB 2; Length 106;
Best Local Similarity 100.0%; Pred. No. 8.9e-06;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GGTTMPGFIADRMQK 15
Db 66 GGTTMPGFIADRMQK 80
```

```
Db 32 GGTTMPGFIADRMQK 46

RESULT 8
Q8JQI2 PRELIMINARY; PRT; 107 AA.
ID Q8JQI2;
AC Q8JQI2;
DT 01-OCT-2002 (TrEMBLrel. 22, Created)
DT 01-OCT-2002 (TrEMBLrel. 22, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE Beta-actin (Fragment).
GN Name=b-act;
OS Anguilla japonica (Japanese eel).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Anguilliformes; Anguillidae;
OC Anguilla.
OX NCBI_TaxID=7937;
RN [1]
RP SEQUENCE FROM N.A.
RA Kurokawa T.;
RT Tissue=Pancreas.
RL Submitted (NOV-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AB074846; BAB96533.1; -.
DR HSSP; P10983; 1D4X.
DR GO; GO:0015629; C:actin cytoskeleton; IEA.
DR GO; GO:0005200; F:structural constituent of cytoskeleton; IEA.
DR InterPro; IPR004000; Actin_like.
DR Pfam; PF00022; Actin; 1.
FT NON_TER 1
FT NON_TER 107
SQ SEQUENCE 107 AA; 12006 MW; 375A7A5C6E6CEA3A CRC64;

Query Match 100.0%; Score 81; DB 2; Length 107;
Best Local Similarity 100.0%; Pred. No. 9e-06;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GGTTMPGFIADRMQK 15
Db 67 GGTTMPGFIADRMQK 81

RESULT 9
P91877 PRELIMINARY; PRT; 109 AA.
ID P91877;
AC P91877;
DT 01-MAY-1997 (TrEMBLrel. 03, Created)
DT 01-MAY-1997 (TrEMBLrel. 03, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE Actin (Fragment).
OS Diprion pini.
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC Neoptera; Endopterygota; Hymenoptera; Tenthredinoidea; Diprionidae;
OC Diprion.
OX NCBI_TaxID=52634;
RN [1]
RP SEQUENCE FROM N.A.
RA Hamelin E., Bigot Y.Y.B., Rouleux F., Renault S., Periquet G.;
RT Submitted (FEB-1997) to the EMBL/GenBank/DBJ databases.
[2]
RP SEQUENCE FROM N.A.
RA Bigot Y.B.;
RL Submitted (SEP-1995) to the EMBL/GenBank/DBJ databases.
DR EMBL; X91507; CAA62807.1; -.
DR HSSP; P10983; 1D4X.
DR GO; GO:0015629; C:actin cytoskeleton; IEA.
DR GO; GO:0005200; F:structural constituent of cytoskeleton; IEA.
DR InterPro; IPR004000; Actin_like.
DR Pfam; PF00022; Actin; 1.
FT NON_TER 1
FT NON_TER 109
SQ SEQUENCE 109 AA; 12036 MW; 15F3926B8D263554 CRC64;

Query Match 100.0%; Score 81; DB 2; Length 109;
```

Best Local Similarity 100.0%; Pred. No. 9.2e-06;  
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGTMYPGIADRMQK 15  
|||||  
Db 54 GGTMYPGIADRMQK 68

## RESULT 10

ID P91881 PRELIMINARY; PRT; 109 AA.  
AC P91881;  
DT 01-MAY-1997 (TrEMBLrel. 03, Created)  
DT 01-MAY-1997 (TrEMBLrel. 03, Last sequence update)  
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)  
DE Actin (Fragment).  
OS Diadromus pulchellus (Parasitic wasp).  
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;  
OC Neoptera; Endopterygota; Hymenoptera; Apocrita; Ichneumonidae;  
OC Ichneumonidae; Ichneumoninae; Diadromus.  
OX NCBI\_TaxID=7420;  
RN [1]  
RP SEQUENCE FROM N.A.  
RA Hamelin E., Bigot Y.Y.B., Rouleux F., Renault S., Periquet G.;  
RL Submitted (FEB-1997) to the EMBL/GenBank/DBJ databases.  
RN [2]  
RP SEQUENCE FROM N.A.  
RA Bigot Y.B.;  
RL Submitted (SEP-1995) to the EMBL/GenBank/DBJ databases.  
DR EMBL; X91506; CAA62806.1; -.  
DR HSSP; P10983; 1D4X.  
DR GO; GO:0015629; C:actin cytoskeleton; IEA.  
DR GO; GO:0005200; E:structural constituent of cytoskeleton; IEA.  
DR InterPro; IPR004000; Actin\_like.  
DR Pfam; PF00022; Actin; 1.  
FT NON\_TER 1  
FT NON\_TER 109  
SQ SEQUENCE 109 AA; 12251 MW; 3C34D292C5AA039A CRC64;

Query Match 100.0%; Score 81; DB 2; Length 109;  
Best Local Similarity 100.0%; Pred. No. 9.2e-06;  
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGTMYPGIADRMQK 15  
|||||  
Db 54 GGTMYPGIADRMQK 68

## RESULT 11

ID O76493 PRELIMINARY; PRT; 110 AA.  
AC O76493;  
DT 01-NOV-1998 (TrEMBLrel. 08, Created)  
DT 01-NOV-1998 (TrEMBLrel. 08, Last sequence update)  
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)  
DE Actin (Fragment).  
OS Bemisia argentifolii (silverleaf whitefly).  
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;  
OC Neoptera; Paraneoptera; Hemiptera; Sternorrhyncha; Aleyrodiformes;  
OC Aleyrodidae; Aleyrodinae; Bemisia.  
OX NCBI\_TaxID=77855;  
RN [1]  
RP SEQUENCE FROM N.A.  
RA He M., Haymer D.S., Lai P.Y.;  
RL Submitted (JUN-1998) to the EMBL/GenBank/DBJ databases.  
DR EMBL; AF071908; AAC24152.1; -.  
DR HSSP; P10983; 1D4X.  
DR GO; GO:0015629; C:actin cytoskeleton; IEA.  
DR GO; GO:0005200; E:structural constituent of cytoskeleton; IEA.  
DR InterPro; IPR004000; Actin\_like.  
DR Pfam; PF00022; Actin; 1.  
FT NON\_TER 1  
FT NON\_TER 110

SQ SEQUENCE 110 AA; 12070 MW; DD2831BC3D4D3EF8 CRC64;

Query Match 100.0%; Score 81; DB 2; Length 110;  
Best Local Similarity 100.0%; Pred. No. 9.3e-06;  
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGTMYPGIADRMQK 15  
|||||  
Db 75 GGTMYPGIADRMQK 89

## RESULT 12

ID Q6R127 PRELIMINARY; PRT; 111 AA.  
AC Q6R127;  
DT 05-JUL-2004 (TrEMBLrel. 27, Created)  
DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)  
DT 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)  
DE Beta-actin (Fragment).  
OS Boophilus microplus (Cattle tick).  
OC Eukaryota; Metazoa; Arthropoda; Chelicerata; Arachnida; Acari;  
OC Parasitiformes; Ixodida; Ixodidae; Boophilus.  
OX NCBI\_TaxID=6941;  
RN [1]  
RP SEQUENCE FROM N.A.  
RA Holmes S.P., Barron P.D., He H., Pietrantonio P.V.;  
RL Submitted (JAN-2004) to the EMBL/GenBank/DBJ databases.  
DR EMBL; AY519367; AAS09968.1; -.  
FT NON\_TER 1  
FT NON\_TER 111  
SQ SEQUENCE 111 AA; 12310 MW; 5639572768E148F5 CRC64;

Query Match 100.0%; Score 81; DB 2; Length 111;  
Best Local Similarity 100.0%; Pred. No. 9.4e-06;  
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGTMYPGIADRMQK 15  
|||||  
Db 69 GGTMYPGIADRMQK 83

## RESULT 13

ID Q6R128 PRELIMINARY; PRT; 111 AA.  
AC Q6R128;  
DT 05-JUL-2004 (TrEMBLrel. 27, Created)  
DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)  
DT 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)  
DE Beta-actin (Fragment).  
OS Boophilus microplus (Cattle tick).  
OC Eukaryota; Metazoa; Arthropoda; Chelicerata; Arachnida; Acari;  
OC Parasitiformes; Ixodida; Ixodidae; Boophilus.  
OX NCBI\_TaxID=6941;  
RN [1]  
RP SEQUENCE FROM N.A.  
RA Holmes S.P., Barron P.D., He H., Pietrantonio P.V.;  
RL Submitted (JAN-2004) to the EMBL/GenBank/DBJ databases.  
DR EMBL; AY519366; AAS09967.1; -.  
FT NON\_TER 1  
FT NON\_TER 111  
SQ SEQUENCE 111 AA; 12294 MW; 40F58D31AA43B48F5 CRC64;

Query Match 100.0%; Score 81; DB 2; Length 111;  
Best Local Similarity 100.0%; Pred. No. 9.4e-06;  
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGTMYPGIADRMQK 15  
|||||  
Db 69 GGTMYPGIADRMQK 83

## RESULT 14

Q6Y260



Search completed: April 8, 2005, 12:03:08  
Job time : 59.8667 secs

```
ID Q6Y260 PRELIMINARY; PRT; 113 AA.
AC Q6Y260;
DT 05-JUL-2004 (TremBLrel. 27, Created)
DT 05-JUL-2004 (TremBLrel. 27, Last sequence update)
DE Alpha-actin protein (Fragment).
OS Pagrus major (Red sea bream) (Chrysophrys major).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;
OC Acanthomorpha; Acanthopterygii; Percomorpha; Perciformes; Percoidae;
OC Sparidae; Pagrus.
OX NCBI_TaxID=143350;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Spleen;
RA Chen S.L., Xu M.Y., Hu S.L., Li L.;
RT "Analysis of immune-relevant genes expressed in red sea bream
   spleen.";
RL Aquaculture 240:115-130(2004).
DR EMBL; AY190676; AAP20152.1; -.
DR GO; GO:0005884; C:actin filament; IEA.
DR GO; GO:0003774; F:motor activity; IEA.
DR GO; GO:0005200; F:structural constituent of cytoskeleton; IEA.
DR InterPro; IPR004001; Actin.
DR PROSITE; PS00432; ACTINS_2; 1.
FT NON_TER 1
SQ SEQUENCE 113 AA; 12550 MW; 4C796A5B5B35269F CRC64;

Query Match 100.0%; Score 81; DB 2; Length 113;
Best Local Similarity 100.0%; Pred. No. 9.5e-06;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GGTMYPGIADRMQK 15
Db 39 GGTMYPGIADRMQK 53

RESULT 15
Q9DF22 PRELIMINARY; PRT; 118 AA.
AC Q9DF22;
DT 01-MAR-2001 (TremBLrel. 16, Created)
DT 01-MAR-2001 (TremBLrel. 16, Last sequence update)
DE Actin (Fragment).
OS Scyliorhinus torazame (Cloudy catshark).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Chondrichthyes;
OC Elasmobranchii; Galeomorphii; Galeoidea; Carchariniiformes;
OC Scyliorhinidae; Scyliorhinus.
OX NCBI_TaxID=75743;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Cartilage;
RA Kim J.T., Kim M.S., Jang E.R., Kim Y.J., Kim K.W.;
RL Submitted (SEP-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF305690; AAG22088.1; -.
DR HSP; P02568; IQ25.
DR GO; GO:0015629; C:actin cytoskeleton; IEA.
DR GO; GO:0005200; F:structural constituent of cytoskeleton; IEA.
DR InterPro; IPR004000; Actin_like.
DR Pfam; PF00022; Actin; 1.
FT NON_TER 1
FT NON_TER 118
SQ SEQUENCE 118 AA; 13380 MW; 9EE39214470C9ACS CRC64;

Query Match 100.0%; Score 81; DB 2; Length 118;
Best Local Similarity 100.0%; Pred. No. 1e-05;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GGTMYPGIADRMQK 15
Db 66 GGTMYPGIADRMQK 80
```

**THIS PAGE BLANK (USPTO)**

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: April 8, 2005, 09:58:56 ; Search time 62.4 Seconds  
(without alignments)  
92.971 Million cell updates/sec

Title: US-09-423-351c-5

Perfect score: 81

Sequence: 1 PRHQGVVMVGQKDS 15

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 2105692 seqs, 386760381 residues

Total number of hits satisfying chosen parameters: 2105692

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Listing first 45 summaries

Database : A\_Geneseq\_16Dec04:\*

- 1: geneseqp1980s:\*
- 2: geneseqp1990s:\*
- 3: geneseqp2000s:\*
- 4: geneseqp2001s:\*
- 5: geneseqp2002s:\*
- 6: geneseqp2003as:\*
- 7: geneseqp2003bs:\*
- 8: geneseqp2004s:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	81	100.0	15	2	Aaw92531 Beta-acti
2	81	100.0	105	5	Abp42469 Human ova
3	81	100.0	157	4	Abb66853 Drosophil
4	81	100.0	201	5	Abp43176 Human ova
5	81	100.0	204	5	Abp43016 Human ova
6	81	100.0	207	4	Abg15288 Novel hum
7	81	100.0	219	8	Abc058422 Human gen
8	81	100.0	310	8	Adr38362 pigA3Cy3G
9	81	100.0	342	7	Adm05323 Human pro
10	81	100.0	374	3	Aay78101 Misgurnus
11	81	100.0	374	3	Abi12985 Human bet
12	81	100.0	375	1	Aap61532 Sequence
13	81	100.0	375	2	Aar50328 Drug resi
14	81	100.0	375	3	Aab15017 Posttrans
15	81	100.0	375	3	Aay94569 Human car
16	81	100.0	375	3	Abi15016 Posttrans
17	81	100.0	375	5	Abb77395 Human act
18	81	100.0	375	6	Abf64271 Angiogene
19	81	100.0	375	6	Abm04830 Rat cytop
20	81	100.0	375	7	Abd85212 Rat actin
21	81	100.0	375	7	Ad61174 Rat Prote
22	81	100.0	375	7	Adf30525 Rat angio
23	81	100.0	375	7	Adi63062 Human apo
24	81	100.0	375	7	Adi62970 Human apo
25	81	100.0	375	7	Adi63040 Human apo

#### ALIGNMENTS

RESULT 1

AAW92531  
ID AAW92531 standard; peptide; 15 AA.

XX AC AAW92531;

DT 26-APR-1999 (first entry)

XX DE Beta-actin reference peptide substrate #5.

KW Peptide substrate; CCT; eukaryotic type II chaperonin complex; cyclin;  
KW binding agent; substrate-binding site; SBS; substrate folding; actin;  
KW tubulin; treatment; cancer; anticancer drug; viral infection; screening;  
KW reduced toxicity.

XX OS Synthetic.

XX FN WO9853322-A1.

XX PD 26-NOV-1998.

XX PF 22-MAY-1998; 98WO-GB001485.

XX PR 23-MAY-1997; 97GB-00010762.

XX PA (CANC-) INST CANCER RES ROYAL CANCER HOSPITAL.

XX PI Willison K, Hynes G, Liou AK;

XX DR WPI; 1999-070162/06.

XX PT Identifying specific binding agents for substrate binding site in CCT  
XX chaperonin complex - also new peptide binding agents and their mimetics,  
XX and peptides containing a specific CCT binding site, used for treating  
XX cancer.

XX PS Disclosure; Fig 10; 97pp; English.

XX CC This invention describes a method which uses the CCT (eukaryotic type II  
XX chaperonin) complex or part of it, for identifying a binding agent that  
XX can occupy a substrate-binding site (SBS) on the CCT complex. By binding  
XX to the CCT complex, the binding agents block an SBS so that biological  
XX activity of the CCT complex is affected, particularly its ability to fold  
XX substrates such as actin, tubulin and cyclin. The binding agents are  
XX useful for treatment of cancer, particularly when used in combination  
XX with an anticancer drug, or viral infections. Nucleic acid fragments are  
XX used to screen for agents, e.g. binding agents that modulate interaction  
XX between the CCT complex and a protein that is to be folded. The binding

ADL13002 Human ste  
ADJ78489 Actin pro  
ADL14103 Human ear  
ADP04899 Sea squir  
ADP12391 Protein e  
ADQ26098 Gamma 1 a  
ABO84772 Human can  
ABm80841 Tumour-as  
ADn23274 Bacterial  
ADs88828 Amino aci  
ADs88825 Amino aci  
ADs88826 Amino aci  
ADs88827 Amino aci  
AAw19799 Gamma-smo  
Abb58162 Drosophil  
Abb61322 Drosophil  
Abb60354 Drosophil  
ABd64853 Drosophil  
ABr62327 Pacific w  
ADn03845 Antipsoi

CC agents may target cells that are actively synthesising tubulin etc.  
 CC (unlike known microtubule-stabilising agents that affect all cells), so  
 CC should have reduced toxicity for normal cells. AM92527-W92541 are  
 CC peptide substrates used in the method of the invention

XX Sequence 15 AA;

Query Match 100.0%; Score 81; DB 2; Length 15;  
 Best Local Similarity 100.0%; Pred. No. 2.8e-07; Indels 0; Gaps 0;  
 Matches 15; Conservative 0; Mismatches 0;

QY 1 PRHQGVWVGGMGQKDS 15  
 |||||  
 Db 1 PRHQGVWVGGMGQKDS 15

RESULT 2  
 ABP42469  
 ID ABP42469 standard; protein; 105 AA.

XX AC ABP42469;

XX DT 22-AUG-2002 (first entry)

XX DE Human ovarian antigen HOCQ94, SEQ ID NO:3601.

XX Human; ovarian antigen; ovary; ovarian; breast; cancer; tumour;  
 KW ovarian cancer; breast cancer; tumour; reproductive system disorder;  
 KW infertility; pregnancy disorder; anovulation; polycystic ovary syndrome;  
 KW PCOS; ovarian cyst; dysmenorrhoea; endocrine disorder; infection;  
 KW inflammatory condition; immune disorder; blood disorder;  
 KW cardiovascular disorder; respiratory disorder; neurological disorder;  
 KW gastrointestinal disorder; urinary system disorder; drug screening;  
 KW gene therapy; chromosome mapping; forensic analysis;  
 KW antibody preparation; cytostatic; immunomodulatory; neuroprotective;  
 KW antiinflammatory; gynaecological; reproductive.

XX Homo sapiens.

XX WO200200677-A1.

XX 03-JAN-2002.

XX 07-JUN-2001; 2001WO-US018569.

XX 07-JUN-2000; 2000US-0209467P.

XX (HUMA-) HUMAN GENOME SCI INC.

XX Birse CE, Rosen CA;

XX WPI; 2002-147878/19.

XX N-PSDB; ABQ55546.

PT Isolated nucleic acid molecules encoding novel ovarian polypeptides,  
 PT useful in the prevention, treatment and diagnosis of cancer (e.g. ovarian  
 PT cancer), immune disorders, cardiovascular disorders and neurological  
 PT diseases.

XX Claim 11; SEQ ID NO 3601; 2922pp; English.

XX The invention relates to 2175 novel human ovarian antigens (ABP41054-  
 CC ABP43228) and to cDNAs encoding them (ABQ54131-ABQ56305), and also  
 CC encompasses polypeptides 90% identical and polynucleotides 95% identical  
 CC to the sequences of the invention. The invention additionally relates to  
 CC recombinant vectors and host cells comprising human ovarian antigen  
 CC polynucleotides, antibodies against human ovarian antigens, and the use  
 CC of ovarian antigen polynucleotides and polypeptides in diagnosing,  
 CC treating, prognosing or preventing various ovary and/or breast-related  
 CC disorders. Such conditions include ovarian cancer and breast cancer, and  
 CC metastatic tumours of ovarian or breast origin, reproductive system  
 CC disorders (e.g., infertility, disorders of pregnancy, anovulation,  
 CC polycystic ovary syndrome, ovarian cysts, and dysmenorrhoea), endocrine

CC disorders, infections (e.g., chlamydia, HIV, toxoplasmosis, and toxic  
 CC shock syndrome), inflammatory conditions (e.g., mastitis, oophoritis and  
 CC vaginitis), immune disorders (e.g., congenital and acquired  
 CC immunodeficiencies, autoimmune oophoritis, systemic lupus erythematosus),  
 CC blood-related disorders (e.g., anaemia), cardiovascular disorders,  
 CC respiratory disorders, neurological disorders, gastrointestinal disorders  
 CC and urinary system disorders. Ovarian antigen polypeptides and  
 CC polynucleotides may also be used in screening for compounds which  
 CC modulate ovarian antigen expression or activity. The polynucleotides may  
 CC further be used for gene therapy, chromosome mapping, in the  
 CC identification of individuals and in forensic analysis, and the  
 CC polypeptides may be used as food additives or to prepare antibodies  
 CC useful in disease diagnosis, drug targeting and phenotyping. The present  
 CC sequence represents a human ovarian antigen of the invention. Note: The  
 CC sequence data for this patent did not form part of the printed  
 CC specification, but was obtained in electronic format directly from WIPO  
 CC at ftp.wipo.int/pub/published\_pct\_sequences

XX SQ Sequence 105 AA;

Query Match 100.0%; Score 81; DB 5; Length 105;  
 Best Local Similarity 100.0%; Pred. No. 2.2e-06;  
 Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 PRHQGVWVGGMGQKDS 15

Db 39 PRHQGVWVGGMGQKDS 53

RESULT 3

ABB66853

ID ABB66853 standard; protein; 157 AA.

XX AC ABB66853;

XX DT 26-MAR-2002 (first entry)

XX Drosophila melanogaster polypeptide SEQ ID NO 27351.

XX Drosophila; developmental biology; cell signalling; insecticide;  
 KW pharmaceutical.

XX Drosophila melanogaster.

XX WO200171042-A2.

XX 27-SEP-2001.

XX 23-MAR-2001; 2001WO-US009231.

XX 23-MAR-2000; 2000US-0191637P.

XX 11-JUL-2000; 2000US-00614150.

XX (PEKE ) PE CORP NY.

XX Venter JC, Adams M, Li PWD, Myers EW;

XX WPI; 2001-656860/75.

XX N-PSDB; ABL10956.

XX New isolated nucleic acid detection reagent for detecting 1000 or more  
 PT genes from Drosophila and for elucidating cell signalling and cell-cell  
 PT interactions.

XX Disclosure; SEQ ID NO 27351; 21pp + Sequence Listing; English.

XX The invention relates to an isolated nucleic acid detection reagent  
 CC capable of detecting 1000 or more genes from Drosophila. The invention is  
 CC useful in developmental biology and in elucidating cell signalling and  
 CC cell-cell interactions in higher eukaryotes for the development of  
 CC insecticides, therapeutics and pharmaceutical drugs. The invention  
 CC discloses genomic DNA sequences (ABL16176-ABL30511), expressed DNA  
 CC sequences (ABL01840-ABL16175) and the encoded proteins (ABB57737-

CC ABB72072). The sequence data for this patent did not form part of the  
 CC printed specification, but was obtained in electronic format directly  
 CC from WIPO at ftp.wipo.int/pub/published\_pct\_sequences  
 XX  
 XX Sequence 157 AA;

Query Match 100.0%; Score 81; DB 4; Length 157;  
 Best Local Similarity 100.0%; Pred. No. 3.5e-06;  
 Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 PRHQGVVMGQKDS 15  
 |||||  
 Db 39 PRHQGVVMGQKDS 53

RESULT 4  
 ABP43176  
 ID ABP43176 standard; protein; 201 AA.  
 AC ABP43176;  
 DT 22-AUG-2002 (first entry)  
 XX Human ovarian antigen HVVCQ49, SEQ ID NO:4308.  
 DE Human; ovarian antigen; ovary; ovarian; breast; cancer; tumour;  
 KW ovarian cancer; breast cancer; tumour; reproductive system disorder;  
 KW infertility; pregnancy disorder; anovulation; polycystic ovary syndrome;  
 KW PCOS; ovarian cyst; dysmenorrhoea; endocrine disorder; infection;  
 KW inflammatory condition; immune disorder; blood disorder;  
 KW cardiovascular disorder; respiratory disorder; neurological disorder;  
 KW gastrointestinal disorder; urinary system disorder; drug screening;  
 KW gene therapy; chromosome mapping; forensic analysis;  
 KW antibody preparation; cytostatic; immunomodulatory; neuroprotective;  
 KW antiinflammatory; gynaecological; reproductive.  
 XX  
 OS Homo sapiens.  
 XX  
 XX WO200200677-A1.  
 PN  
 XX  
 XX 03-JAN-2002.  
 PD  
 XX  
 XX 07-JUN-2001; 2001WO-US018569.  
 XX  
 XX 07-JUN-2000; 2000US-0209467P.  
 PR  
 XX  
 XX (HUMA-) HUMAN GENOME SCI INC.  
 PA  
 XX  
 XX Birse CE, Rosen CA;  
 PI  
 XX  
 XX WPI; 2002-147878/19.  
 DR  
 XX  
 XX N-PSDB; ABQ56253.  
 DR  
 XX  
 XX Isolated nucleic acid molecules encoding novel ovarian polypeptides,  
 PT useful in the prevention, treatment and diagnosis of cancer (e.g. ovarian  
 PT cancer), immune disorders, cardiovascular disorders and neurological  
 PT diseases.  
 XX  
 XX Claim 11; SEQ ID NO 4308; 2922pp; English.  
 PS  
 XX  
 XX The invention relates to 2175 novel human ovarian antigens (ABP41054-  
 CC ABP43228) and to cDNAs encoding them (ABQ54131-ABQ56305), and also  
 CC encompasses polypeptides 90% identical and polynucleotides 95% identical  
 CC to the sequences of the invention. The invention additionally relates to  
 CC recombinant vectors and host cells comprising human ovarian antigen  
 CC polynucleotides, antibodies against human ovarian antigens, and the use  
 CC of ovarian antigen polynucleotides and polypeptides in diagnosing,  
 CC treating, prognosing or preventing various ovary and/or breast-related  
 CC disorders. Such conditions include ovarian cancer and breast cancer, and  
 CC metastatic tumours of ovarian or breast origin, reproductive system  
 CC disorders (e.g., infertility, disorders of pregnancy, anovulation,  
 CC polycystic ovary syndrome, ovarian cysts, and dysmenorrhoea), endocrine  
 CC disorders, infections (e.g., chlamydia, HIV, toxoplasmosis, and toxic

CC shock syndrome), inflammatory conditions (e.g., mastitis, oophoritis and  
 CC vaginitis), immune disorders (e.g., congenital and acquired  
 CC immunodeficiencies, autoimmune oophoritis, systemic lupus erythematosus),  
 CC blood-related disorders (e.g., anaemia), cardiovascular disorders,  
 CC respiratory disorders, neurological disorders, gastrointestinal disorders  
 CC and urinary system disorders. Ovarian antigen polypeptides and  
 CC polynucleotides may also be used in screening for compounds which  
 CC modulate ovarian antigen expression or activity. The polynucleotides may  
 CC further be used for gene therapy, chromosome mapping, in the  
 CC identification of individuals and in forensic analysis, and the  
 CC polypeptides may be used as food additives or to prepare antibodies  
 CC useful in disease diagnosis, drug targeting and phenotyping. The present  
 CC sequence represents a human ovarian antigen of the invention. Note: The  
 CC sequence data for this patent did not form part of the printed  
 CC specification, but was obtained in electronic format directly from WIPO  
 CC at ftp.wipo.int/pub/published\_pct\_sequences  
 XX  
 XX Sequence 201 AA;

Query Match 100.0%; Score 81; DB 5; Length 201;  
 Best Local Similarity 100.0%; Pred. No. 4.5e-06;  
 Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 PRHQGVVMGQKDS 15  
 |||||  
 Db 55 PRHQGVVMGQKDS 69

RESULT 5  
 ABP43016  
 ID ABP43016 standard; protein; 204 AA.  
 XX ABP43016;  
 AC ABP43016;  
 XX  
 DT 22-AUG-2002 (first entry)  
 XX  
 XX Human ovarian antigen HSPSB24, SEQ ID NO:4148.  
 DE Human; ovarian antigen; ovary; ovarian; breast; cancer; tumour;  
 KW ovarian cancer; breast cancer; tumour; reproductive system disorder;  
 KW infertility; pregnancy disorder; anovulation; polycystic ovary syndrome;  
 KW PCOS; ovarian cyst; dysmenorrhoea; endocrine disorder; infection;  
 KW inflammatory condition; immune disorder; blood disorder;  
 KW cardiovascular disorder; respiratory disorder; neurological disorder;  
 KW gastrointestinal disorder; urinary system disorder; drug screening;  
 KW gene therapy; chromosome mapping; forensic analysis;  
 KW antibody preparation; cytostatic; immunomodulatory; neuroprotective;  
 KW antiinflammatory; gynaecological; reproductive.  
 XX  
 OS Homo sapiens.  
 XX  
 XX WO200200677-A1.  
 PN  
 XX  
 XX 03-JAN-2002.  
 PD  
 XX  
 XX 07-JUN-2001; 2001WO-US018569.  
 XX  
 XX 07-JUN-2000; 2000US-0209467P.  
 PR  
 XX  
 XX (HUMA-) HUMAN GENOME SCI INC.  
 PA  
 XX  
 XX Birse CE, Rosen CA;  
 PI  
 XX  
 XX WPI; 2002-147878/19.  
 DR  
 XX  
 XX N-PSDB; ABQ56093.  
 DR  
 XX  
 XX Isolated nucleic acid molecules encoding novel ovarian polypeptides,  
 PT useful in the prevention, treatment and diagnosis of cancer (e.g. ovarian  
 PT cancer), immune disorders, cardiovascular disorders and neurological  
 PT diseases.  
 XX  
 XX Claim 11; SEQ ID NO 4148; 2922pp; English.  
 PS  
 XX

CC The invention relates to 2175 novel human ovarian antigens (ABP41054-  
 CC ABP43228) and to cDNAs encoding them (AB054131-AB056305), and also  
 CC encompasses polypeptides 90% identical and polynucleotides 95% identical  
 CC to the sequences of the invention. The invention additionally relates to  
 CC recombinant vectors and host cells comprising human ovarian antigen  
 CC polynucleotides, antibodies against human ovarian antigens, and the use  
 CC of ovarian antigen polynucleotides and polypeptides in diagnosing,  
 CC treating, prognosing or preventing various ovary and/or breast-related  
 CC disorders. Such conditions include ovarian cancer and breast cancer, and  
 CC metastatic tumours of ovarian or breast origin, reproductive system  
 CC disorders (e.g., infertility, disorders of pregnancy, anovulation,  
 CC polycystic ovary syndrome, ovarian cysts, and dysmenorrhea), endocrine  
 CC disorders, infections (e.g., chlamydia, HIV, toxoplasmosis, and toxic  
 CC shock syndrome), inflammatory conditions (e.g., mastitis, oophoritis and  
 CC vaginitis), immune disorders (e.g., congenital and acquired  
 CC immunodeficiencies, autoimmune oophoritis, systemic lupus erythematosus),  
 CC blood-related disorders (e.g., anaemia), cardiovascular disorders,  
 CC respiratory disorders, neurological disorders, gastrointestinal disorders,  
 CC and urinary system disorders. Ovarian antigen polypeptides and  
 CC polynucleotides may also be used in screening for compounds which  
 CC modulate ovarian antigen expression or activity. The polynucleotides may  
 CC further be used for gene therapy, chromosome mapping, in the  
 CC identification of individuals and in forensic analysis, and the  
 CC polypeptides may be used as food additives or to prepare antibodies  
 CC useful in disease diagnosis, drug targeting and phenotyping. The present  
 CC sequence represents a human ovarian antigen of the invention. Note: The  
 CC sequence data for this patent did not form part of the printed  
 CC specification, but was obtained in electronic format directly from WIPO  
 CC at ftp.wipo.int/pub/published\_pct\_sequences

XX SQ Sequence 204 AA;

Query Match 100.0%; Score 81; DB 5; Length 204;  
 Best Local Similarity 100.0%; Pred. No. 4.6e-06;  
 Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 1 PRHQGVVMVGMGQKDS 15  
 |||||  
 Db 58 PRHQGVVMVGMGQKDS 72

RESULT 6

ABG15288  
 ID ABG15288 standard; protein; 207 AA.

XX AC ABG15288;

XX DT 18-FEB-2002 (first entry)

XX DE Novel human diagnostic protein #15279.

XX DE Human; chromosome mapping; gene mapping; gene therapy; forensic;  
 KW food supplement; medical imaging; diagnostic; genetic disorder.

XX OS Homo sapiens.

XX PN WO200175067-A2.

XX PD 11-OCT-2001.

XX PF 30-MAR-2001; 2001WO-US008631.

XX PR 31-MAR-2000; 2000US-00540217.

XX PR 23-AUG-2000; 2000US-00649167.

XX XX (HYSE-) HYSEQ INC.

XX PI Drmanac RT, Liu C, Tang YT;

XX DR WPI; 2001-639362/73.

XX DR N-PSDB; AAS79475.

XX PT New isolated polynucleotide and encoded polypeptides, useful in

PT diagnostics, forensics, gene mapping, identification of mutations  
 PT responsible for genetic disorders or other traits and to assess  
 PT biodiversity.

XX PS Claim 20; SEQ ID NO 45647; 103pp; English.

XX CC The invention relates to isolated polynucleotide (I) and polypeptide (II)  
 CC sequences. (I) is useful as hybridisation probes, polymerase chain  
 CC reaction (PCR) primers, oligomers, and for chromosome and gene mapping,  
 CC and in recombinant production of (II). The polynucleotides are also used  
 CC in diagnostics as expressed sequence tags for identifying expressed  
 CC genes. (I) is useful in gene therapy techniques to restore normal  
 CC activity of (II) or to treat disease states involving (II). (II) is  
 CC useful for generating antibodies against it, detecting or quantitating a  
 CC polypeptide in tissue, as molecular weight markers and as a food  
 CC supplement. (II) and its binding partners are useful in medical imaging  
 CC of sites expressing (II). (I) and (II) are useful for treating disorders  
 CC involving aberrant protein expression or biological activities. The  
 CC polypeptide and polynucleotide sequences have applications in  
 CC diagnostics, forensics, gene mapping, identification of mutations  
 CC responsible for genetic disorders or other traits to assess biodiversity  
 CC and to produce other types of data and products dependent on DNA and  
 CC amino acid sequences. ABG00010-ABG30377 represent novel human diagnostic  
 CC amino acid sequences of the invention. Note: The sequence data for this  
 CC patent did not appear in the printed specification, but was obtained in  
 CC electronic format directly from WIPO at  
 CC ftp.wipo.int/pub/published\_pct\_sequences

XX SQ Sequence 207 AA;

Query Match 100.0%; Score 81; DB 4; Length 207;

Best Local Similarity 100.0%; Pred. No. 4.7e-06;

Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 1 PRHQGVVMVGMGQKDS 15

|||||  
 Db 65 PRHQGVVMVGMGQKDS 79

RESULT 7

ABO58422

ID ABO58422 standard; protein; 219 AA.

XX AC ABO58422;

XX DT 29-JUL-2004 (first entry)

XX DE Human genome derived single exon protein #4656.

XX DE Human; gene expression; single exon probe; microarray;  
 KW alternative splicing event; genomic alteration.

XX OS Homo sapiens.

XX PN US2003194704-A1.

XX PD 16-OCT-2003.

XX PF 03-APR-2002; 2002US-00029386.

XX PF 03-APR-2002; 2002US-00029386.

XX XX (PENN/) PENN S G.

XX PA (RANK/) RANK D R.

XX PA (HANZ/) HANZEL D K.

XX XX Penn SG, Rank DR, Hanzel DK;

XX XX WPI; 2004-119264/12.

XX PT New human genome-derived single exon nucleic acid probes useful for human  
 PT gene expression analysis, for identifying or characterizing alternative  
 PT splicing events, for assessing genomic alterations or as tools for



PI Seki N, Yoshikawa T, Otsuka M, Nagahari K, Masuho Y;  
 XX WPI; 2003-723558/69.  
 DR N-PSDB; ADM02880.  
 XX  
 PT New polynucleotides and polypeptides are useful in gene therapy, for  
 PT developing a diagnostic marker or medicines for regulating their  
 PT expression and activity, or as a target of gene therapy.  
 XX  
 PS Claim 1; SEQ ID NO 4008; 305pp; English.  
 XX  
 CC The invention relates to a novel human polynucleotide and the encoded  
 CC polypeptide. A polynucleotide of the invention may have a use in gene  
 CC therapy. An oligonucleotide of the invention ADM06202-ADM06773 is useful  
 CC as a primer for synthesizing the polynucleotide or as a probe for  
 CC detecting the polynucleotide. The polynucleotides ADM01316-ADM03758 are  
 CC useful in gene therapy, for developing a diagnostic marker or medicines  
 CC for regulating their expression and activity, or as a target of gene  
 CC therapy. The proteins ADM03759-ADM06201 encoded by the polynucleotides  
 CC are useful as pharmaceutical agents. The present sequence represents a  
 CC protein sequence of the invention.  
 XX  
 SQ Sequence 342 AA;  
 Query Match 100.0%; Score 81; DB 7; Length 342;  
 Best Local Similarity 100.0%; Pred. No. 8e-06;  
 Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 Qy 1 PRHQGVVMVGQKDS 15  
 Db 40 PRHQGVVMVGQKDS 54  
 RESULT 10  
 AAY78101  
 ID AAY78101 standard; protein; 374 AA.  
 XX  
 AC AAY78101;  
 XX  
 DT 27-APR-2000 (first entry)  
 XX  
 DE Misgurnus mizolepis beta actin.  
 XX  
 KW Growth hormone; beta actin; controlling site; fish; breeding.  
 XX  
 OS Misgurnus mizolepis.  
 XX  
 PN JP2000004889-A.  
 XX  
 PD 11-JAN-2000.  
 XX  
 PF 01-JUN-1999; 99JP-00154346.  
 XX  
 PR 01-JUN-1999; 98KR-00020255.  
 PR 23-JUL-1998; 98KR-00029701.  
 XX  
 PA (KIMD/) KIM D S.  
 PA (KIMC/) KIM C G.  
 PA (NOHJ/) NOH J G.  
 PA (CHOK/) CHO K N.  
 PA (NAMY/) NAM Y G.  
 XX  
 PI Cho KN, Kim CG, Kim DS, Nam YG, Noh JG;  
 XX  
 XX WPI; 2000-140126/13.  
 DR N-PSDB; AAZ88225.  
 XX  
 PT A Misgurnus mizolepis growth hormone expression vector - containing DNA  
 PT encoding a beta-actin gene.  
 XX  
 PS Claim 1; Page 16-20; 30pp; Japanese.  
 XX  
 CC The present invention describes a Misgurnus mizolepis growth hormone

CC expression vector. Also described is the DNA sequence given in AAZ88225  
 CC which contains the beta actin gene and beta actin gene controlling site  
 CC of Misgurnus mizolepis. Also described are: (1) an expression vector  
 CC containing a growth hormone gene and the beta actin gene controlling site  
 CC of Misgurnus mizolepis; and (2) a method for the preparation of a high  
 CC growth transformed Misgurnus mizolepis including a step of finely  
 CC injecting the above expression vector to fertilised eggs of Misgurnus  
 CC mizolepis and hatching them. The vector can be used for the expression of  
 CC useful genes in a fish. The method can reduce the breeding period of  
 CC Misgurnus mizolepis. The present sequence represents the Misgurnus  
 CC mizolepis beta actin amino acid sequence  
 XX  
 SQ Sequence 374 AA;  
 Query Match 100.0%; Score 81; DB 3; Length 374;  
 Best Local Similarity 100.0%; Pred. No. 8.8e-06;  
 Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 Qy 1 PRHQGVVMVGQKDS 15  
 Db 38 PRHQGVVMVGQKDS 52  
 RESULT 11  
 AAB12985  
 ID AAB12985 standard; protein; 374 AA.  
 XX  
 AC AAB12985;  
 XX  
 DT 29-NOV-2000 (first entry)  
 XX  
 DE Human beta-actin protein sequence.  
 XX  
 KW Beta actin; post translational modification; sickle cell anaemia;  
 KW irreversibly sickled cell; ISC; treatment.  
 XX  
 OS Homo sapiens.  
 XX  
 PN US6087398-A.  
 XX  
 PD 11-JUL-2000.  
 XX  
 PF 01-MAR-1996; 96US-00609236.  
 XX  
 PR 14-AUG-1995; 95US-0002288P.  
 XX  
 PA (SALA-) SOUTH ALABAMA MEDICAL SCI FOUND.  
 XX  
 PI Goodman SR;  
 XX  
 PD WPI; 2000-498210/44.  
 XX  
 PT Treatment for sickle cell anemia comprises administering a reducing agent  
 PT to inhibit and reverse sickled cell formation in blood.  
 XX  
 PS Example 17; Fig 5B; 53pp; English.  
 XX  
 CC This invention provides a treatment for sickle cell anaemia, which  
 CC comprises administering a reducing agent. The treatment inhibits the  
 CC formation of irreversible sickled cells (ISC) and reverses ISC formation  
 CC in the blood. The present sequence represents the human beta-actin  
 CC protein. The sequence is used in the invention to demonstrate that a  
 CC disulphide bridge is formed between cysteines 284 and 373 in ISC beta-  
 CC actin as a post-translational modification. The reducing agent used in  
 CC the treatment acts to correct this post-translational modification  
 XX  
 SQ Sequence 374 AA;  
 Query Match 100.0%; Score 81; DB 3; Length 374;  
 Best Local Similarity 100.0%; Pred. No. 8.8e-06;  
 Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 Qy 1 PRHQGVVMVGQKDS 15



```
Db      37 PRHQGVVMVGQKDS 51
|||||
RESULT 12
AAP61532
ID AAP61532 standard; protein; 375 AA.
XX
AC AAP61532;
XX
XX 25-MAR-2003 (revised)
DT 13-SEP-1991 (first entry)
XX
DE Sequence of beta-actin.
XX
KW Actin fibre; actin film; recombinant beta-actin.
XX
OS Homo sapiens.
XX
PN EP174608-A.
XX
PD 19-MAR-1986.
XX
PF 05-SEP-1985; 85EP-00111225.
XX
PR 13-SEP-1984; 84US-00650958.
XX
PA (STRD ) UNIV LELAND STANFORD JUNIOR.
XX
PI Leavitt JC, Kedes LH, Gunning PW;
XX
DR WPI; 1986-077015/12.
DR N-PSDB; AAN60172.
XX
PT Beta-actin gene and regulatory elements - used for expression of
PT polypeptide(s) in mammalian host cells.
XX
PS Example; Page 23-24; 32pp; English.
XX
CC In the example, a beta-actin expression vector providing the beta- actin
CC promoter region, a polylinker and a polyadenylation signal was
CC constructed where the expression construct was present on a vector having
CC a bacterial origin of replication, as well as a marker for selection in a
CC mammalian host. (Updated on 25-MAR-2003 to correct PA field.)
XX
SQ Sequence 375 AA;
Query Match 100.0%; Score 81; DB 1; Length 375;
Best Local Similarity 100.0%; Pred. No. 8.8e-06;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 PRHQGVVMVGQKDS 15
Db 38 PRHQGVVMVGQKDS 52
|||||
RESULT 13
AAR50328
ID AAR50328 standard; protein; 375 AA.
XX
AC AAR50328;
XX
DT 12-OCT-1994 (first entry)
XX
DE Drug resistant structural protein.
XX
XX Drug resistant; structural gene; expression vector; selective; marker;
KW transformation; thymidine-kinase deleted cell.
XX
XX Homo sapiens.
XX
PN JP06038773-A.
XX
PD 15-FEB-1994.
XX
PF 31-JAN-1992; 92JP-00045939.
XX
PR 31-JAN-1992; 92JP-00045939.
XX
PA (TOYM ) TOYOBO KK.
XX
DR WPI; 1994-094836/12.
DR N-PSDB; AAQ44861.
XX
PT Expression vector contg. drug resistant marker gene - useful for
PT transformation of thymidine kinase-deleted human cells for high levels of
PT foreign protein prodn.
XX
PS Disclosure; Page 5-7; 7pp; Japanese.
XX
CC This sequence is encoded by a drug resistant structural gene which may be
CC used in the expression vector of the invention. This gene is used as the
CC selective marker in the vector. The resulting vector may be used to
CC transform a thymidine-kinase deleted cell allowing introduction of a
CC foreign structural gene. The transformed cell may be used to produce
CC large amounts of useful protein
XX
SQ Sequence 375 AA;
Query Match 100.0%; Score 81; DB 2; Length 375;
Best Local Similarity 100.0%; Pred. No. 8.8e-06;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 PRHQGVVMVGQKDS 15
Db 38 PRHQGVVMVGQKDS 52
|||||
RESULT 14
AAB15017
ID AAB15017 standard; protein; 375 AA.
XX
AC AAB15017;
XX
DT 07-DEC-2000 (first entry)
XX
DE Posttranslationally modified human cardiac actin mutant E361G.
XX
KW Cardiac actin; ACTC; human; gene therapy; IDC; chromosome 15q14;
KW idiopathic dilated cardiomyopathy; mutant; mutein.
XX
OS Homo sapiens.
XX
XX Synthetic.
FH Key Location/Qualifiers
FT Misc-difference 361
FT /note= "Wild-type Glu substituted by Gly"
XX
XX US6063576-A.
XX
PD 16-MAY-2000.
XX
PF 29-JUN-1998; 98US-00106217.
XX
PR 29-JUN-1998; 98US-00106217.
XX
PA (UTAH ) UNIV UTAH RES FOUND.
XX
PI Olson TM, Keating MT;
XX
DR WPI; 2000-375488/32.
DR N-PSDB; AAA73739.
XX
XX New cardiac actin gene comprising histidine to arginine or glycine to
XX glutamic acid substitution, useful in the diagnosis of diseases
XX associated with the mutation, specifically idiopathic dilated
```

CC one in exon 5 and one in exon 6. The mutations are G to A in codon 312  
 CC (Arg312His) and A to G in codon 361 (Glu361Gly). The ACTC protein may be  
 CC used in a variety of methods for drug screening and for rational drug  
 CC design. The ACTC gene may be used to treat IDC by gene therapy. Analysis  
 CC of the ACTC gene provides early identification of subjects likely to  
 CC develop or who already have IDC  
 XX  
 SQ Sequence 375 AA;  
 Query Match 100.0%; Score 81; DB 3; Length 375;  
 Best Local Similarity 100.0%; Pred. No. 8.8e-06;  
 Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 PRHQGVVVGMGQKDS 15  
 |||||  
 Db 38 PRHQGVVVGMGQKDS 52  
 |||||  
 Search completed: April 8, 2005, 10:50:43  
 Job time : 62.4 secs

PT cardiomyopathy.  
 XX Claim 3; Page: 36pp; English.  
 PS  
 XX The present invention relates to human cardiac actin (ACTC). Genotypic  
 CC analyses show that ACTC is linked to idiopathic dilated cardiomyopathy  
 CC (IDC). The ACTC gene maps to chromosome 15q14. Six PCR primer pairs  
 CC (AAA49020 to AAA49031) have been developed to allow analysis of the six  
 CC exons of ACTC. Studies of two families with IDC showed the presence of  
 CC two mutations, one in exon 5 and one in exon 6. The present sequence is a  
 CC mutant of posttranslationally modified ACTC without the initial two  
 CC residues. The mutation is a Glu to Gly mutation at position 361. The ACTC  
 CC protein may be used in a variety of methods for drug screening and for  
 CC rational drug design. The ACTC gene may be used to treat IDC by gene  
 CC therapy. Analysis of the ACTC gene provides early identification of  
 CC subjects likely to develop or who already have IDC. Note: The present  
 CC sequence is not shown in the specification but is derived from the ACTC  
 CC sequence described in AAY94569  
 XX  
 SQ Sequence 375 AA;  
 Query Match 100.0%; Score 81; DB 3; Length 375;  
 Best Local Similarity 100.0%; Pred. No. 8.8e-06;  
 Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 PRHQGVVVGMGQKDS 15  
 |||||  
 Db 38 PRHQGVVVGMGQKDS 52  
 |||||

RESULT 15  
 AAY94569  
 ID AAY94569 standard; protein; 375 AA.  
 XX  
 AC AAY94569;  
 XX  
 XX 07-DEC-2000 (first entry)  
 DT  
 XX Human cardiac actin protein after posttranslational modification.  
 DE  
 XX Cardiac actin; ACTC; human; gene therapy; IDC; 15q14;  
 KW idiopathic dilated cardiomyopathy.  
 XX  
 XX Homo sapiens.  
 OS  
 XX US6063576-A.  
 FN  
 XX 16-MAY-2000.  
 PD  
 XX 29-JUN-1998; 98US-00106217.  
 PF  
 XX 29-JUN-1998; 98US-00106217.  
 PR  
 XX (UTAH ) UNIV UTAH RES FOUND.  
 PA  
 XX Olson TM, Keating MT;  
 PI  
 XX WPI; 2000-375488/32.  
 DR  
 XX N-PSDB; AAA49032.  
 DR  
 XX New cardiac actin gene comprising histidine to arginine or glycine to  
 PT glutamic acid substitution, useful in the diagnosis of diseases  
 PT associated with the mutation, specifically idiopathic dilated  
 PT cardiomyopathy.  
 XX  
 PS Claim 3; Col 61-62; 36pp; English.  
 PS  
 XX The present sequence is human cardiac actin (ACTC) protein, without the  
 CC initial Met-Cys which is posttranslationally removed. Genotypic analyses  
 CC show that ACTC is linked to idiopathic dilated cardiomyopathy (IDC). The  
 CC ACTC gene maps to chromosome 15q14. Six PCR primer pairs (AAA49020 to  
 CC AAA49031) have been developed to allow analysis of the six exons of ACTC.  
 CC Studies of two families with IDC showed the presence of two mutations,

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: April 8, 2005, 09:58:57 ; Search time 17.5333 Seconds  
(without alignments)  
63.863 Million cell updates/sec

Title: US-09-423-351C-5

Perfect score: 81

Sequence: 1 PRHQVVMGMQKDS 15

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 513545 seqs, 74649064 residues

Total number of hits satisfying chosen parameters: 513545

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Listing first 45 summaries

Database : Issued Patents AA.\*

- 1: /cgn2\_6/ptodata/1/iaa/5A\_COMB.pep.\*
- 2: /cgn2\_6/ptodata/1/iaa/5B\_COMB.pep.\*
- 3: /cgn2\_6/ptodata/1/iaa/6A\_COMB.pep.\*
- 4: /cgn2\_6/ptodata/1/iaa/6B\_COMB.pep.\*
- 5: /cgn2\_6/ptodata/1/iaa/PTUS\_COMB.pep.\*
- 6: /cgn2\_6/ptodata/1/iaa/backfiles1.pep.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	81	100.0	374	3	US-08-609-236-6
2	81	100.0	374	3	US-09-306-446C-2
3	81	100.0	375	2	US-08-494-151-14
4	81	100.0	375	3	US-09-106-217-16
5	81	100.0	375	4	US-09-976-594-731
6	81	100.0	376	1	US-08-588-113-2
7	81	100.0	376	4	US-09-949-016-6100
8	81	100.0	377	3	US-09-106-217-2
9	81	100.0	377	4	US-09-919-172-33
10	81	100.0	377	4	US-09-917-254-53
11	81	100.0	386	4	US-09-949-016-7721
12	81	100.0	399	4	US-09-949-016-9424
13	81	100.0	402	4	US-09-949-016-10757
14	81	100.0	404	4	US-09-949-016-11313
15	80	98.8	377	4	US-09-248-796A-14109
16	77	95.1	20	3	US-08-505-250-17
17	77	95.1	20	3	US-08-505-250-17
18	76	93.8	71	4	US-09-621-976-6521
19	74	91.4	375	3	US-09-171-337A-7
20	74	91.4	375	3	US-09-171-337A-8
21	74	91.4	375	4	US-09-631-022-7
22	74	91.4	375	4	US-09-631-022-8
23	56	69.1	80	3	US-09-306-446C-4
24	48	59.3	371	1	US-08-261-206A-77
25	45	55.6	397	4	US-09-902-540-9949
26	43	53.1	72	4	US-09-252-991A-28861
27	43	53.1	843	4	US-09-252-991A-32609

28	41	50.6	543	4	US-09-252-991A-18055	Sequence 18055, A
29	40	49.4	211	4	US-09-252-991A-18464	Sequence 18464, A
30	40	49.4	832	4	US-09-540-236-3056	Sequence 3056, Ap
31	40	49.4	862	4	US-09-543-681A-6315	Sequence 6315, Ap
32	39	48.1	334	4	US-09-949-016-11006	Sequence 11006, A
33	39	48.1	338	4	US-09-252-991A-22180	Sequence 22180, A
34	39	48.1	453	4	US-09-252-991A-31634	Sequence 31634, A
35	39	48.1	817	4	US-09-252-991A-27609	Sequence 27609, A
36	38	46.9	120	4	US-09-765-815-3	Sequence 3, Appli
37	38	46.9	245	4	US-09-902-540-13736	Sequence 13736, A
38	38	46.9	266	4	US-09-270-767-4486	Sequence 4486, A
39	38	46.9	516	4	US-09-252-991A-17933	Sequence 17933, A
40	38	46.9	545	4	US-09-538-092-1297	Sequence 1297, A
41	38	46.9	652	1	US-08-765-081-6	Sequence 6, Appli
42	38	46.9	652	3	US-09-098-082-6	Sequence 6, Appli
43	38	46.9	957	4	US-09-252-991A-30672	Sequence 30672, A
44	37	45.7	62	4	US-09-248-796A-22522	Sequence 22522, A
45	37	45.7	66	4	US-09-543-681A-7224	Sequence 7224, Ap

## ALIGNMENTS

RESULT 1  
US-08-609-236-6  
; Sequence 6, Application US/08609236  
; Patent No. 6087398  
; GENERAL INFORMATION:  
; APPLICANT: Steven R. Goodman  
; TITLE OF INVENTION: NO. 6087398el Sickie Cell Anemia Treatment  
; NUMBER OF SEQUENCES: 6  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: McGregor & Adler, LLP  
; STREET: 8011 Candle Lane  
; CITY: Houston  
; STATE: Texas  
; COUNTRY: USA  
; ZIP: 77071  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: Apple Macintosh  
; OPERATING SYSTEM: Macintosh  
; SOFTWARE: Microsoft Word for Macintosh  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/609,236  
; FILING DATE: March 1, 1996  
; CLASSIFICATION: 514  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 60/002,288  
; FILING DATE: August 14, 1995  
; CLASSIFICATION: 514  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Benjamin Aaron Adler, Ph.D.  
; REGISTRATION NUMBER: 35,423  
; REFERENCE/DOCKET NUMBER: D5807  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 713-777-2321  
; TELEFAX: 713-777-6908  
; INFORMATION FOR SEQ ID NO: 6:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 374  
; TYPE: Amino acid  
; STRANDEDNESS:  
; TOPOLOGY: Linear  
; MOLECULE TYPE:  
; DESCRIPTION: Protein  
; HYPOTHETICAL: No  
; ANTI-SENSE: No  
; FRAGMENT TYPE:  
; ORIGINAL SOURCE:  
; STRAIN:  
; INDIVIDUAL ISOLATE:  
; DEVELOPMENTAL STAGE:

```
;
; TISSUE TYPE:
; CELL TYPE:
; CELL LINE:
US-08-609-236-6

Query Match      100.0%; Score 81; DB 3; Length 374;
Best Local Similarity 100.0%; Pred. No. 1.9e-06;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 PRHQGVVMVGQKDS 15
Db 37 PRHQGVVMVGQKDS 51

RESULT 2
US-09-306-446C-2
; Sequence 2, Application US/09306446C
; Patent No. 6372959
; GENERAL INFORMATION:
; APPLICANT: KIM, Dong Soo
; APPLICANT: KIM, Chul Geun
; APPLICANT: NAM, Yoon Kwon
; APPLICANT: NOH, Jae Koo
; APPLICANT: CHO, Kyou Nam
; TITLE OF INVENTION: EXPRESSION VECTOR OF MUD LOACH GROWTH HORMONE GENE
; FILE REFERENCE: P06344US/BA5
; CURRENT FILING DATE: 1999-05-06
; PRIOR APPLICATION NUMBER: US/09/306,446C
; PRIOR FILING DATE: KR 98/20255
; NUMBER OF SEQ ID NOS: 33
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 2
; LENGTH: 374
; TYPE: PRT
; ORGANISM: Misgurnus mizolepus
US-09-306-446C-2

Query Match      100.0%; Score 81; DB 3; Length 374;
Best Local Similarity 100.0%; Pred. No. 1.9e-06;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 PRHQGVVMVGQKDS 15
Db 38 PRHQGVVMVGQKDS 52

RESULT 3
US-08-494-151-14
; Sequence 14, Application US/08494151
; Patent No. 5840528
; GENERAL INFORMATION:
; APPLICANT: Van Ooyen, Albert Johannes Joseph
; TITLE OF INVENTION: Transformation of Phatfia rhodozyma
; NUMBER OF SEQUENCES: 14
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Morrison & Foerster
; STREET: 2000 Pennsylvania Ave. N.W., Suite 5500
; CITY: Washington, D.C.
; COUNTRY: USA
; ZIP: 20006-1812
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; FILING DATE: 23-JUN-1995
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Murashige, Kate H.
; REGISTRATION NUMBER: 29,959

; REFERENCE/DOCKET NUMBER: 24615-20039.01
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202) 887-1500
; TELEFAX: (202) 887-0763
; TELEX: 90-4030
; INFORMATION FOR SEQ ID NO: 14:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 375 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-494-151-14

Query Match      100.0%; Score 81; DB 2; Length 375;
Best Local Similarity 100.0%; Pred. No. 1.9e-06;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 PRHQGVVMVGQKDS 15
Db 38 PRHQGVVMVGQKDS 52

RESULT 4
US-09-106-217-16
; Sequence 16, Application US/09106217
; Patent No. 6063576
; GENERAL INFORMATION:
; APPLICANT: Keating, Mark T.
; APPLICANT: Olson, Timothy M.
; TITLE OF INVENTION: Actin Mutations in Dilated
; TITLE OF INVENTION: Cardiomyopathy, a Heritable Form of Heart Failure
; NUMBER OF SEQUENCES: 18
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Rothwell, Figg, Ernst & Kurz, P.C.
; STREET: 555 Thirteenth Street, N.W., Suite 701 East
; CITY: Washington
; STATE: DC
; COUNTRY: U.S.A.
; ZIP: 20004
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; FILING DATE:
; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
; NAME: Saxe, Stephen A.
; REGISTRATION NUMBER: 38,609
; REFERENCE/DOCKET NUMBER: 2323-125
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 202-783-6040
; TELEFAX: 202-783-6031
; INFORMATION FOR SEQ ID NO: 16:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 375 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-09-106-217-16

Query Match      100.0%; Score 81; DB 3; Length 375;
Best Local Similarity 100.0%; Pred. No. 1.9e-06;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 PRHQGVVMVGQKDS 15
Db 38 PRHQGVVMVGQKDS 52
```

```

RESULT 6
US-08-588-113-2
; Sequence 2, Application US/08588113
; Patent No. 5710003
; GENERAL INFORMATION:
; APPLICANT: McHugh, Kirk M.
; TITLE OF INVENTION: DIAGNOSTIC TEST FOR DETERMINING
; TITLE OF INVENTION: MALIGNANCY OF SMOOTH MUSCLE TUMORS
; NUMBER OF SEQUENCES: 4
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Woodcock Washburn Kurtz Mackiewicz &
; ADDRESSEE: No. 5710003irs
; STREET: One Liberty Place - 46th Floor
; CITY: Philadelphia
; STATE: PA
; COUNTRY: USA
; ZIP: 19103
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC Compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/588,113
; FILING DATE:
; CLASSIFICATION: 436
; ATTORNEY/AGENT INFORMATION:
; NAME: Ralph, Rebecca L.
; REGISTRATION NUMBER: 35,152
; REFERENCE/DOCKET NUMBER: TJU-1652
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 215-568-3100
; TELEFAX: 215-568-3439
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 376 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-588-113-2

```

; REGISTRATION NUMBER: 38,609  
; REFERENCE/DOCKET NUMBER: 2323-125  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 202-783-6040  
; TELEFAX: 202-783-6031  
; INFORMATION FOR SEQ ID NO: 2:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 377 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
US-09-106-217-2

Query Match 100.0%; Score 81; DB 3; Length 377;  
Best Local Similarity 100.0%; Pred. No. 2e-06;  
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 PRHQGVVMVGMGQKDS 15  
Db 40 PRHQGVVMVGMGQKDS 54

## RESULT 9

US-09-919-172-33  
; Sequence 33, Application US/09919172  
; Patent No. 6673545  
; GENERAL INFORMATION:  
; APPLICANT: Faris, Mary  
; APPLICANT: Turner, Christopher M.  
; TITLE OF INVENTION: PROSTATE CANCER MARKERS  
; FILE REFERENCE: PA-0036 US  
; CURRENT APPLICATION NUMBER: US/09/919,172  
; CURRENT FILING DATE: 2001-07-30  
; PRIOR APPLICATION NUMBER: 60/222,469  
; PRIOR FILING DATE: 2000-07-28  
; NUMBER OF SEQ ID NOS: 102  
; SOFTWARE: PERL Program  
; SEQ ID NO 33  
; LENGTH: 377  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
; FEATURE:  
; NAME/KEY: misc feature  
; OTHER INFORMATION: Incyte ID No. 6673545 1709118CD1  
US-09-919-172-33

Query Match 100.0%; Score 81; DB 4; Length 377;  
Best Local Similarity 100.0%; Pred. No. 2e-06;  
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 PRHQGVVMVGMGQKDS 15  
Db 40 PRHQGVVMVGMGQKDS 54

## RESULT 10

US-09-917-254-53  
; Sequence 53, Application US/09917254  
; Patent No. 6703204  
; GENERAL INFORMATION:  
; APPLICANT: Mutter, George  
; APPLICANT: Baak, Jan  
; TITLE OF INVENTION: Prognostic Classification of Breast Cancer  
; FILE REFERENCE: B0801/7224(JRV)  
; CURRENT APPLICATION NUMBER: US/09/917,254  
; CURRENT FILING DATE: 2001-07-27  
; PRIOR APPLICATION NUMBER: US 60/222,093  
; PRIOR FILING DATE: 2000-07-28  
; NUMBER OF SEQ ID NOS: 102  
; SOFTWARE: PatentIn version 3.0  
; SEQ ID NO 53  
; LENGTH: 377  
; TYPE: PRT

; ORGANISM: Homo Sapiens  
US-09-917-254-53

Query Match 100.0%; Score 81; DB 4; Length 377;  
Best Local Similarity 100.0%; Pred. No. 2e-06;  
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 PRHQGVVMVGMGQKDS 15  
Db 40 PRHQGVVMVGMGQKDS 54

## RESULT 11

US-09-949-016-7721  
; Sequence 7721, Application US/09949016  
; Patent No. 6812339  
; GENERAL INFORMATION:  
; APPLICANT: VENTER, J. Craig et al.  
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED  
; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF  
; FILE REFERENCE: CL001307  
; CURRENT APPLICATION NUMBER: US/09/949,016  
; CURRENT FILING DATE: 2000-04-14  
; PRIOR APPLICATION NUMBER: 60/241,755  
; PRIOR FILING DATE: 2000-10-20  
; PRIOR APPLICATION NUMBER: 60/237,768  
; PRIOR FILING DATE: 2000-10-03  
; PRIOR APPLICATION NUMBER: 60/231,498  
; PRIOR FILING DATE: 2000-09-08  
; NUMBER OF SEQ ID NOS: 207012  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 7721  
; LENGTH: 386  
; TYPE: PRT  
; ORGANISM: Human  
US-09-949-016-7721

Query Match 100.0%; Score 81; DB 4; Length 386;  
Best Local Similarity 100.0%; Pred. No. 2e-06;  
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 PRHQGVVMVGMGQKDS 15  
Db 49 PRHQGVVMVGMGQKDS 63

## RESULT 12

US-09-949-016-9424  
; Sequence 9424, Application US/09949016  
; Patent No. 6812339  
; GENERAL INFORMATION:  
; APPLICANT: VENTER, J. Craig et al.  
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED  
; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF  
; FILE REFERENCE: CL001307  
; CURRENT APPLICATION NUMBER: US/09/949,016  
; CURRENT FILING DATE: 2000-04-14  
; PRIOR APPLICATION NUMBER: 60/241,755  
; PRIOR FILING DATE: 2000-10-20  
; PRIOR APPLICATION NUMBER: 60/237,768  
; PRIOR FILING DATE: 2000-10-03  
; PRIOR APPLICATION NUMBER: 60/231,498  
; PRIOR FILING DATE: 2000-09-08  
; NUMBER OF SEQ ID NOS: 207012  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 9424  
; LENGTH: 399  
; TYPE: PRT  
; ORGANISM: Human  
US-09-949-016-9424

Query Match 100.0%; Score 81; DB 4; Length 399;  
Best Local Similarity 100.0%; Pred. No. 2.1e-06;

Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 PRHQGVVMVGMGQKDS 15  
| | | | | | | | | | | | | | | |  
Db 62 PRHQGVVMVGMGQKDS 76

## RESULT 13

US-09-949-016-10757  
; Sequence 10757, Application US/09949016  
; Patent No. 6812339  
; GENERAL INFORMATION:  
; APPLICANT: VENTER, J. Craig et al.  
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED  
; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF  
; FILE REFERENCE: CL001307  
; CURRENT APPLICATION NUMBER: US/09/949,016  
; CURRENT FILING DATE: 2000-04-14  
; PRIOR FILING DATE: 60/241,755  
; PRIOR FILING DATE: 2000-10-20  
; PRIOR APPLICATION NUMBER: 60/237,768  
; PRIOR FILING DATE: 2000-10-03  
; PRIOR APPLICATION NUMBER: 60/231,498  
; PRIOR FILING DATE: 2000-09-08  
; NUMBER OF SEQ ID NOS: 207012  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 10757  
; LENGTH: 402  
; TYPE: PRT  
; ORGANISM: Human  
US-09-949-016-10757

Query Match 100.0%; Score 81; DB 4; Length 402;  
Best Local Similarity 100.0%; Pred. No. 2.1e-06;  
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 PRHQGVVMVGMGQKDS 15  
| | | | | | | | | | | | | | | |  
Db 65 PRHQGVVMVGMGQKDS 79

## RESULT 14

US-09-949-016-11313  
; Sequence 11313, Application US/09949016  
; Patent No. 6812339  
; GENERAL INFORMATION:  
; APPLICANT: VENTER, J. Craig et al.  
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED  
; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF  
; FILE REFERENCE: CL001307  
; CURRENT APPLICATION NUMBER: US/09/949,016  
; CURRENT FILING DATE: 2000-04-14  
; PRIOR APPLICATION NUMBER: 60/241,755  
; PRIOR FILING DATE: 2000-10-20  
; PRIOR APPLICATION NUMBER: 60/237,768  
; PRIOR FILING DATE: 2000-10-03  
; PRIOR APPLICATION NUMBER: 60/231,498  
; PRIOR FILING DATE: 2000-09-08  
; NUMBER OF SEQ ID NOS: 207012  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 11313  
; LENGTH: 404  
; TYPE: PRT  
; ORGANISM: Human  
US-09-949-016-11313

Query Match 100.0%; Score 81; DB 4; Length 404;  
Best Local Similarity 100.0%; Pred. No. 2.1e-06;  
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 PRHQGVVMVGMGQKDS 15  
| | | | | | | | | | | | | | | |  
Db 67 PRHQGVVMVGMGQKDS 81

## RESULT 15

US-09-248-796A-14109  
; Sequence 14109, Application US/09248796A  
; Patent No. 6747137  
; GENERAL INFORMATION:  
; APPLICANT: Keith Weinstock et al  
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO CANDIDA ALBICAN  
; AND THERAPEUTICS  
; FILE REFERENCE: 107196.132  
; CURRENT APPLICATION NUMBER: US/09/248,796A  
; CURRENT FILING DATE: 1999-02-12  
; PRIOR APPLICATION NUMBER: US 60/074,725  
; PRIOR FILING DATE: 1998-02-13  
; PRIOR APPLICATION NUMBER: US 60/096,409  
; PRIOR FILING DATE: 1998-08-13  
; NUMBER OF SEQ ID NOS: 28208  
; SEQ ID NO 14109  
; LENGTH: 377  
; TYPE: PRT  
; ORGANISM: Candida albicans  
US-09-248-796A-14109

Query Match 98.8%; Score 80; DB 4; Length 377;  
Best Local Similarity 93.3%; Pred. No. 3e-06;  
Matches 14; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 PRHQGVVMVGMGQKDS 15  
| | | | | | | | | | | | | | | |  
Db 40 PRHQGVVMVGMGQKDS 54

Search completed: April 8, 2005, 12:07:35  
Job time : 17.5333 secs

**THIS PAGE BLANK (USPTO)**



GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: April 8, 2005, 10:53:18 ; Search time 42.6667 Seconds  
(without alignments)  
116.718 Million cell updates/sec

Title: US-09-423-351C-5

Perfect score: 81

Sequence: 1 PRHQGVVMVGQKDS 15

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1418010 seqs, 331997259 residues

Total number of hits satisfying chosen parameters: 1418010

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Listing first 45 summaries

Database : Published Applications AA:\*

- 1: /cgn2\_6/ptodata/1/pubpaa/US07\_PUBCOMB.pep.\*
- 2: /cgn2\_6/ptodata/1/pubpaa/PCT\_NEW\_PUB.pep.\*
- 3: /cgn2\_6/ptodata/1/pubpaa/US05\_NEW\_PUB.pep.\*
- 4: /cgn2\_6/ptodata/1/pubpaa/US06\_PUBCOMB.pep.\*
- 5: /cgn2\_6/ptodata/1/pubpaa/US07\_NEW\_PUB.pep.\*
- 6: /cgn2\_6/ptodata/1/pubpaa/PCTUS\_PUBCOMB.pep.\*
- 7: /cgn2\_6/ptodata/1/pubpaa/US08\_NEW\_PUB.pep.\*
- 8: /cgn2\_6/ptodata/1/pubpaa/US08\_PUBCOMB.pep.\*
- 9: /cgn2\_6/ptodata/1/pubpaa/US09A\_PUBCOMB.pep.\*
- 10: /cgn2\_6/ptodata/1/pubpaa/US09B\_PUBCOMB.pep.\*
- 11: /cgn2\_6/ptodata/1/pubpaa/US09C\_PUBCOMB.pep.\*
- 12: /cgn2\_6/ptodata/1/pubpaa/US09\_NEW\_PUB.pep.\*
- 13: /cgn2\_6/ptodata/1/pubpaa/US10A\_PUBCOMB.pep.\*
- 14: /cgn2\_6/ptodata/1/pubpaa/US10B\_PUBCOMB.pep.\*
- 15: /cgn2\_6/ptodata/1/pubpaa/US10C\_PUBCOMB.pep.\*
- 16: /cgn2\_6/ptodata/1/pubpaa/US10D\_PUBCOMB.pep.\*
- 17: /cgn2\_6/ptodata/1/pubpaa/US10\_NEW\_PUB.pep.\*
- 18: /cgn2\_6/ptodata/1/pubpaa/US11\_NEW\_PUB.pep.\*
- 19: /cgn2\_6/ptodata/1/pubpaa/US60\_NEW\_PUB.pep.\*
- 20: /cgn2\_6/ptodata/1/pubpaa/US60\_PUBCOMB.pep.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	81	100.0	105	US-10-264-049-3601	Sequence 3601, Ap
2	81	100.0	142	US-10-424-599-184471	Sequence 184471, Ap
3	81	100.0	201	US-10-264-049-4308	Sequence 4308, Ap
4	81	100.0	204	US-10-284-049-4148	Sequence 4148, Ap
5	81	100.0	219	US-10-002-631C-56	Sequence 56, Appl
6	81	100.0	219	US-10-029-386-32056	Sequence 32056, A
7	81	100.0	342	US-10-108-260A-4008	Sequence 4008, Ap
8	81	100.0	371	US-10-322-281-68	Sequence 68, Appl
9	81	100.0	375	US-10-205-194-93	Sequence 93, Appl
10	81	100.0	375	US-10-316-253-88	Sequence 88, Appl
11	81	100.0	375	US-10-369-493-5927	Sequence 5927, Ap
12	81	100.0	375	US-10-205-331-94	Sequence 94, Appl
13	81	100.0	375	US-10-260-708-82	Sequence 82, Appl

14	81	100.0	376	15	US-10-341-434-101	Sequence 101, App
15	81	100.0	377	9	US-09-919-172-33	Sequence 33, Appl
16	81	100.0	377	15	US-10-236-031B-64	Sequence 64, Appl
17	81	100.0	377	16	US-10-333-680-4	Sequence 4, Appl
18	81	100.0	393	15	US-10-092-900A-310	Sequence 310, App
19	81	100.0	398	16	US-10-322-281-63	Sequence 63, Appl
20	81	100.0	413	9	US-09-925-301-1436	Sequence 1436, Ap
21	80	98.8	375	15	US-10-369-493-1786	Sequence 1786, Ap
22	80	98.8	375	16	US-10-477-369-1	Sequence 1, Appl
23	75	92.6	375	15	US-10-369-493-2436	Sequence 2436, Ap
24	72	88.9	60	15	US-10-424-599-161246	Sequence 161246, A
25	72	88.9	73	15	US-10-424-599-202833	Sequence 202833, A
26	72	88.9	108	16	US-10-767-701-35226	Sequence 35226, A
27	72	88.9	135	16	US-10-767-701-47318	Sequence 47318, A
28	72	88.9	136	15	US-10-424-599-280811	Sequence 280811, A
29	72	88.9	197	15	US-10-424-599-280810	Sequence 280810, A
30	72	88.9	208	15	US-10-424-599-282905	Sequence 282905, A
31	72	88.9	230	16	US-10-767-701-42770	Sequence 42770, A
32	72	88.9	248	15	US-10-424-599-162684	Sequence 162684, A
33	72	88.9	347	15	US-10-424-599-283333	Sequence 283333, A
34	72	88.9	376	16	US-10-437-963-179917	Sequence 179917, A
35	72	88.9	376	16	US-10-437-963-202420	Sequence 202420, A
36	72	88.9	376	16	US-10-767-701-45677	Sequence 45677, A
37	72	88.9	376	16	US-10-767-701-46090	Sequence 46090, A
38	72	88.9	377	14	US-10-338-777-52	Sequence 52, Appl
39	72	88.9	377	15	US-10-424-599-162009	Sequence 162009, A
40	72	88.9	377	15	US-10-424-599-162685	Sequence 162685, A
41	72	88.9	377	15	US-10-424-599-162686	Sequence 162686, A
42	72	88.9	377	15	US-10-424-599-222816	Sequence 222816, A
43	72	88.9	377	15	US-10-424-599-223492	Sequence 223492, A
44	72	88.9	377	15	US-10-424-599-283336	Sequence 283336, A
45	72	88.9	377	15	US-10-425-114-66124	Sequence 66124, A

#### ALIGNMENTS

RESULT 1  
US-10-264-049-3601  
; Sequence 3601, Application US/10264049  
; Publication No. US20040005579A1  
; GENERAL INFORMATION:  
; APPLICANT: Birse et al.  
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies  
; FILE REFERENCE: P413P1  
; CURRENT APPLICATION NUMBER: US/10/264,049  
; PRIOR FILING DATE: 2002-10-04  
; PRIOR APPLICATION NUMBER: PCT/US01/18569  
; PRIOR FILING DATE: 2001-06-07  
; PRIOR APPLICATION NUMBER: US 60/209,467  
; PRIOR FILING DATE: 2000-06-07  
; NUMBER OF SEQ ID NOS: 4360  
; SOFTWARE: Patent In Ver. 3.1  
; SEQ ID NO 3601  
; LENGTH: 105  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
; FEATURE:  
; NAME/KEY: MISC FEATURE  
; LOCATION: (103)  
; OTHER INFORMATION: Xaa equals any of the twenty naturally occurring L-amino acids  
US-10-264-049-3601

Query Match 100.0%; Score 81; DB 15; Length 105;  
Best Local Similarity 100.0%; Pred. No. 2.2e-06;  
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 PRHQGVVMVGQKDS 15  
Db 39 PRHQGVVMVGQKDS 53

RESULT 2

```
US-10-424-599-184471
; Sequence 184471, Application US/10424599
; Publication No. US20040031072A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa Thomas J
; APPLICANT: Kovalic David K
; APPLICANT: Zhou Yihua
; APPLICANT: Cao Yongwei
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With
; FILE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53223)B
; CURRENT APPLICATION NUMBER: US/10/424,599
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 285684
; SEQ ID NO 184471
; LENGTH: 142
; TYPE: PRT
; ORGANISM: Glycine max
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT3847_137594C.1.pep
US-10-424-599-184471

Query Match      100.0%; Score 81; DB 15; Length 142;
Best Local Similarity 100.0%; Pred. No. 3.1e-06;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 PRHOGVMVGMGQKDS 15
Db 47 PRHOGVMVGMGQKDS 61

RESULT 3
US-10-264-049-4308
; Sequence 4308, Application US/10264049
; Publication No. US20040005579A1
; GENERAL INFORMATION:
; APPLICANT: Birse et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
; FILE REFERENCE: PA133P1
; CURRENT APPLICATION NUMBER: US/10/264,049
; CURRENT FILING DATE: 2002-10-04
; PRIOR APPLICATION NUMBER: PCT/US01/18569
; PRIOR FILING DATE: 2001-06-07
; PRIOR APPLICATION NUMBER: US 60/209,467
; PRIOR FILING DATE: 2000-06-07
; NUMBER OF SEQ ID NOS: 4360
; SOFTWARE: PatentIn Ver. 3.1
; SEQ ID NO 4308
; LENGTH: 201
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: MISC FEATURE
; LOCATION: (135)
; OTHER INFORMATION: Xaa equals any of the twenty naturally occurring L-amino acids
; NAME/KEY: MISC FEATURE
; LOCATION: (136)
; OTHER INFORMATION: Xaa equals any of the twenty naturally occurring L-amino acids
; NAME/KEY: MISC FEATURE
; LOCATION: (140)
; OTHER INFORMATION: Xaa equals any of the twenty naturally occurring L-amino acids
; FEATURE:
; NAME/KEY: MISC FEATURE
; LOCATION: (143)
; OTHER INFORMATION: Xaa equals any of the twenty naturally occurring L-amino acids
; FEATURE:
; NAME/KEY: MISC FEATURE
; LOCATION: (144)
; OTHER INFORMATION: Xaa equals any of the twenty naturally occurring L-amino acids
; FEATURE:
; NAME/KEY: MISC FEATURE
```

```
; LOCATION: (145)
; OTHER INFORMATION: Xaa equals any of the twenty naturally occurring L-amino acids
; FEATURE:
; NAME/KEY: MISC FEATURE
; LOCATION: (151)
; OTHER INFORMATION: Xaa equals any of the twenty naturally occurring L-amino acids
; FEATURE:
; NAME/KEY: MISC FEATURE
; LOCATION: (159)
; OTHER INFORMATION: Xaa equals any of the twenty naturally occurring L-amino acids
; NAME/KEY: MISC FEATURE
; LOCATION: (172)
; OTHER INFORMATION: Xaa equals any of the twenty naturally occurring L-amino acids
; FEATURE:
; NAME/KEY: MISC FEATURE
; LOCATION: (174)
; OTHER INFORMATION: Xaa equals any of the twenty naturally occurring L-amino acids
; NAME/KEY: MISC FEATURE
; LOCATION: (180)
; OTHER INFORMATION: Xaa equals any of the twenty naturally occurring L-amino acids
US-10-264-049-4308

Query Match      100.0%; Score 81; DB 15; Length 201;
Best Local Similarity 100.0%; Pred. No. 4.4e-06;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 PRHOGVMVGMGQKDS 15
Db 55 PRHOGVMVGMGQKDS 69

RESULT 4
US-10-264-049-4148
; Sequence 4148, Application US/10264049
; Publication No. US20040005579A1
; GENERAL INFORMATION:
; APPLICANT: Birse et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
; FILE REFERENCE: PA133P1
; CURRENT APPLICATION NUMBER: US/10/264,049
; CURRENT FILING DATE: 2002-10-04
; PRIOR APPLICATION NUMBER: PCT/US01/18569
; PRIOR FILING DATE: 2001-06-07
; PRIOR APPLICATION NUMBER: US 60/209,467
; PRIOR FILING DATE: 2000-06-07
; NUMBER OF SEQ ID NOS: 4360
; SOFTWARE: PatentIn Ver. 3.1
; SEQ ID NO 4148
; LENGTH: 204
; TYPE: PRT
; ORGANISM: Homo sapiens
; NAME/KEY: MISC FEATURE
; LOCATION: (135)
; OTHER INFORMATION: Xaa equals any of the twenty naturally occurring L-amino acids
; NAME/KEY: MISC FEATURE
; LOCATION: (136)
; OTHER INFORMATION: Xaa equals any of the twenty naturally occurring L-amino acids
; NAME/KEY: MISC FEATURE
; LOCATION: (140)
; OTHER INFORMATION: Xaa equals any of the twenty naturally occurring L-amino acids
; FEATURE:
; NAME/KEY: MISC FEATURE
; LOCATION: (143)
; OTHER INFORMATION: Xaa equals any of the twenty naturally occurring L-amino acids
; FEATURE:
; NAME/KEY: MISC FEATURE
; LOCATION: (144)
; OTHER INFORMATION: Xaa equals any of the twenty naturally occurring L-amino acids
; FEATURE:
; NAME/KEY: MISC FEATURE
; LOCATION: (145)
; OTHER INFORMATION: Xaa equals any of the twenty naturally occurring L-amino acids
; FEATURE:
; NAME/KEY: MISC FEATURE
; LOCATION: (151)
; OTHER INFORMATION: Xaa equals any of the twenty naturally occurring L-amino acids
; FEATURE:
; NAME/KEY: MISC FEATURE
; LOCATION: (159)
; OTHER INFORMATION: Xaa equals any of the twenty naturally occurring L-amino acids
; NAME/KEY: MISC FEATURE
; LOCATION: (172)
; OTHER INFORMATION: Xaa equals any of the twenty naturally occurring L-amino acids
; FEATURE:
; NAME/KEY: MISC FEATURE
; LOCATION: (174)
; OTHER INFORMATION: Xaa equals any of the twenty naturally occurring L-amino acids
; NAME/KEY: MISC FEATURE
; LOCATION: (180)
; OTHER INFORMATION: Xaa equals any of the twenty naturally occurring L-amino acids
US-10-264-049-4308

Query Match      100.0%; Score 81; DB 15; Length 204;
Best Local Similarity 100.0%; Pred. No. 4.5e-06;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 PRHOGVMVGMGQKDS 15
Db 58 PRHOGVMVGMGQKDS 72

RESULT 5
US-10-002-631C-56
; Sequence 56, Application US/10002631C
; Publication No. US20030157486A1
; GENERAL INFORMATION:
; APPLICANT: Graff, Jonathon M.
; APPLICANT: Muenster, Matthew
; TITLE OF INVENTION: METHODS TO IDENTIFY SIGNAL SEQUENCES
; FILE REFERENCE: A34943 090495.0243
```

```

; CURRENT APPLICATION NUMBER: US/10/002,631C
; CURRENT FILING DATE: 2001-10-31
; PRIOR APPLICATION NUMBER: 60/300,309
; PRIOR FILING DATE: 2001-06-21
; NUMBER OF SEQ ID NOS: 324
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 56
; LENGTH: 219
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-002-631C-56

Query Match      100.0%; Score 81; DB 14; Length 219;
Best Local Similarity 100.0%; Pred. No. 4.8e-06;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      1 PRHQGVVMVGMGQKDS 15
Db      65 PRHQGVVMVGMGQKDS 79

RESULT 6
US-10-029-386-32056
; Sequence 32056, Application US/10029386
; Publication No. US20030194704A1
; GENERAL INFORMATION:
; APPLICANT: Penn, Sharron G.
; APPLICANT: Rank, David R.
; APPLICANT: Hanzel, David K.
; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR
; FILE REFERENCE: AEOMICA-X-2
; CURRENT APPLICATION NUMBER: US/10/029,386
; CURRENT FILING DATE: 2001-12-20
; NUMBER OF SEQ ID NOS: 34288
; SOFTWARE: Annomax Sequence Listing Engine vers. 1.1
; SEQ ID NO 32056
; LENGTH: 219
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: MAP TO AC026717.3
; OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 16
; OTHER INFORMATION: EXPRESSED IN HELA, SIGNAL = 23
; OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 6.8
; OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 20
; OTHER INFORMATION: EXPRESSED IN HEART, SIGNAL = 9.8
; OTHER INFORMATION: SWISSPROT HIT: P29751, EVALUATE 3.00e-94
US-10-029-386-32056

Query Match      100.0%; Score 81; DB 14; Length 219;
Best Local Similarity 100.0%; Pred. No. 4.8e-06;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      1 PRHQGVVMVGMGQKDS 15
Db      69 PRHQGVVMVGMGQKDS 83

RESULT 7
US-10-108-260A-4008
; Sequence 4008, Application US/10108260A
; Publication No. US20040005560A1
; GENERAL INFORMATION:
; APPLICANT: HELIX RESEARCH INSTITUTE
; TITLE OF INVENTION: No. US20040005560A1el full length cDNA
; FILE REFERENCE: H1-A0106
; CURRENT APPLICATION NUMBER: US/10/108,260A
; CURRENT FILING DATE: 2002-03-27
; NUMBER OF SEQ ID NOS: 5458
; SOFTWARE: PatentIn ver. 2.1
; SEQ ID NO 4008
; LENGTH: 342

```

```

; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-108-260A-4008

Query Match      100.0%; Score 81; DB 15; Length 342;
Best Local Similarity 100.0%; Pred. No. 7.7e-06;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      1 PRHQGVVMVGMGQKDS 15
Db      40 PRHQGVVMVGMGQKDS 54

RESULT 8
US-10-322-281-68
; Sequence 68, Application US/10322281
; Publication No. US20040126762A1
; GENERAL INFORMATION:
; APPLICANT: David W. Morris
; APPLICANT: Marc S. Malandro
; TITLE OF INVENTION: Novel Compositions and Methods in Cancer
; FILE REFERENCE: 529452001000
; CURRENT APPLICATION NUMBER: US/10/322,281
; CURRENT FILING DATE: 2002-12-17
; NUMBER OF SEQ ID NOS: 866
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 68
; LENGTH: 371
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-322-281-68

Query Match      100.0%; Score 81; DB 16; Length 371;
Best Local Similarity 100.0%; Pred. No. 8.4e-06;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      1 PRHQGVVMVGMGQKDS 15
Db      38 PRHQGVVMVGMGQKDS 52

RESULT 9
US-10-205-194-93
; Sequence 93, Application US/10205194
; Publication No. US20030134301A1
; GENERAL INFORMATION:
; APPLICANT: Warner-Lambert Company
; APPLICANT: Lee, Kevin
; APPLICANT: Dixon, Alistair
; APPLICANT: Brooksbank, Robert
; APPLICANT: Pinnock, Robert
; TITLE OF INVENTION: Identification and Use of Molecules Implicated in Pain
; FILE REFERENCE: WL-A-018201
; CURRENT APPLICATION NUMBER: US/10/205,194
; CURRENT FILING DATE: 2000-07-24
; PRIOR APPLICATION NUMBER: GB 0118354.0
; PRIOR FILING DATE: 2001-07-27
; NUMBER OF SEQ ID NOS: 177
; SOFTWARE: PatentIn ver. 2.1
; SEQ ID NO 93
; LENGTH: 375
; TYPE: PRT
; ORGANISM: Rattus rattus
; FEATURE:
; OTHER INFORMATION: Cytoplasmic gamma isoform of actin
US-10-205-194-93

Query Match      100.0%; Score 81; DB 14; Length 375;
Best Local Similarity 100.0%; Pred. No. 8.5e-06;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      1 PRHQGVVMVGMGQKDS 15

```

```
Db          38 PRHQGVVMVGMGQKDS 52

RESULT 10
US-10-316-253-88
; Sequence 88, Application US/10316253
; Publication No. US20030162706A1
; GENERAL INFORMATION:
; APPLICANT: The Procter & Gamble Company
; APPLICANT: Peters, Kevin
; APPLICANT: Thompson, Larry
; APPLICANT: Wang, Feng
; APPLICANT: Greis, Kenneth
; TITLE OF INVENTION: Angiogenesis Modulating Proteins
; FILE REFERENCE: 8865M
; CURRENT APPLICATION NUMBER: US/10/316,253
; CURRENT FILING DATE: 2002-12-10
; PRIOR APPLICATION NUMBER: US 60/355,295
; PRIOR FILING DATE: 2002-02-08
; NUMBER OF SEQ ID NOS: 308
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 88
; LENGTH: 375
; TYPE: PRT
; ORGANISM: Rattus norvegicus
US-10-316-253-88

Query Match          100.0%; Score 81; DB 14; Length 375;
Best Local Similarity 100.0%; Pred. No. 8.5e-06;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 PRHQGVVMVGMGQKDS 15
      |||||
Db      38 PRHQGVVMVGMGQKDS 52

RESULT 11
US-10-369-493-5927
; Sequence 5927, Application US/10369493
; Publication No. US20030233675A1
; GENERAL INFORMATION:
; APPLICANT: Cao, Yongwei
; APPLICANT: Hinkle, Gregory J.
; APPLICANT: Slater, Steven C.
; APPLICANT: Goldman, Barry S.
; APPLICANT: Chen, Xianfeng
; TITLE OF INVENTION: EXPRESSION OF MICROBIAL PROTEINS IN PLANTS FOR PRODUCTION OF
; FILE REFERENCE: 38-10(52052)B
; CURRENT APPLICATION NUMBER: US/10/369,493
; CURRENT FILING DATE: 2003-02-28
; PRIOR APPLICATION NUMBER: US 60/360,039
; PRIOR FILING DATE: 2002-02-21
; NUMBER OF SEQ ID NOS: 47374
; SEQ ID NO 5927
; LENGTH: 375
; TYPE: PRT
; ORGANISM: Caenorhabditis elegans
US-10-369-493-5927

Query Match          100.0%; Score 81; DB 15; Length 375;
Best Local Similarity 100.0%; Pred. No. 8.5e-06;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 PRHQGVVMVGMGQKDS 15
      |||||
Db      38 PRHQGVVMVGMGQKDS 52

RESULT 12
US-10-205-331-94
; Sequence 94, Application US/10205331
; Publication No. US20040058326A1
```

```
; GENERAL INFORMATION:
; APPLICANT: Warner-Lambert Company
; APPLICANT: Lee, Kevin
; APPLICANT: Dixon, Alistair
; APPLICANT: Brooksbank, Robert
; APPLICANT: Pincock, Robert
; TITLE OF INVENTION: Identification and Use of Molecules Implicated in Pain
; FILE REFERENCE: WL-A-018199
; CURRENT APPLICATION NUMBER: US/10/205,331
; CURRENT FILING DATE: 2002-07-24
; PRIOR APPLICATION NUMBER: GB 0118354.0
; PRIOR FILING DATE: 2001-07-27
; NUMBER OF SEQ ID NOS: 117
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 94
; LENGTH: 375
; TYPE: PRT
; ORGANISM: Rattus norvegicus
; FEATURE:
; OTHER INFORMATION: Cytoplasmic beta-actin
US-10-205-331-94

Query Match          100.0%; Score 81; DB 15; Length 375;
Best Local Similarity 100.0%; Pred. No. 8.5e-06;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 PRHQGVVMVGMGQKDS 15
      |||||
Db      38 PRHQGVVMVGMGQKDS 52

RESULT 13
US-10-260-708-82
; Sequence 82, Application US/10260708
; Publication No. US20040063101A1
; GENERAL INFORMATION:
; APPLICANT: Scanlan, Matthew
; APPLICANT: Lee, Sang-Yull
; APPLICANT: Old, Lloyd
; TITLE OF INVENTION: Human Sarcoma-Associated Antigens
; FILE REFERENCE: L00461/70138
; CURRENT APPLICATION NUMBER: US/10/260,708
; CURRENT FILING DATE: 2002-09-30
; NUMBER OF SEQ ID NOS: 96
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 82
; LENGTH: 375
; TYPE: PRT
; ORGANISM: homo sapiens
US-10-260-708-82

Query Match          100.0%; Score 81; DB 15; Length 375;
Best Local Similarity 100.0%; Pred. No. 8.5e-06;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 PRHQGVVMVGMGQKDS 15
      |||||
Db      38 PRHQGVVMVGMGQKDS 52

RESULT 14
US-10-341-434-101
; Sequence 101, Application US/10341434
; Publication No. US20030215835A1
; GENERAL INFORMATION:
; APPLICANT: Origene Technologies
; TITLE OF INVENTION: Differentially Regulated Prostate Cancer Genes
; FILE REFERENCE: 9U 204 205 R1
; CURRENT APPLICATION NUMBER: US/10/341,434
; CURRENT FILING DATE: 2003-07-18
; PRIOR APPLICATION NUMBER: US 60/348,164
; PRIOR FILING DATE: 2002-01-15
; PRIOR APPLICATION NUMBER: US 60/348,119
```

; PRIOR FILING DATE: 2002-01-15  
 ; NUMBER OF SEQ ID NOS: 238  
 ; SOFTWARE: PatentIn version 3.1  
 ; SEQ ID NO 101  
 ; LENGTH: 376  
 ; TYPE: PRT  
 ; ORGANISM: Homo sapiens  
 US-10-341-434-101

Query Match 100.0%; Score 81; DB 15; Length 376;  
 Best Local Similarity 100.0%; Pred. No. 8.5e-06;  
 Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 PRHQGVVVGQKDS 15  
 Db 39 PRHQGVVVGQKDS 53

RESULT 15

US-09-919-172-33  
 ; Sequence 33, Application US/09919172  
 ; Patent No. US20020119463A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Farris, Mary  
 ; TITLE OF INVENTION: PROSTATE CANCER MARKERS  
 ; FILE REFERENCE: PA-0036 US  
 ; CURRENT APPLICATION NUMBER: US/09/919,172  
 ; CURRENT FILING DATE: 2001-07-30  
 ; PRIOR APPLICATION NUMBER: 60/222,469  
 ; PRIOR FILING DATE: 2000-07-28  
 ; NUMBER OF SEQ ID NOS: 102  
 ; SOFTWARE: PERL Program  
 ; SEQ ID NO 33  
 ; LENGTH: 377  
 ; TYPE: PRT  
 ; ORGANISM: Homo sapiens  
 ; FEATURE:  
 ; NAME/KEY: misc feature  
 ; OTHER INFORMATION: Incyte ID No. US20020119463A1 1709118CD1  
 US-09-919-172-33

Query Match 100.0%; Score 81; DB 9; Length 377;  
 Best Local Similarity 100.0%; Pred. No. 8.5e-06;  
 Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 PRHQGVVVGQKDS 15  
 Db 40 PRHQGVVVGQKDS 54

Search completed: April 8, 2005, 12:50:59  
 Job time : 43.6667 secs

**THIS PAGE BLANK (USPTO)**

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: April 8, 2005, 09:58:56 ; Search time 9.33333 Seconds  
(without alignment)  
154.634 Million cell updates/sec

Title: US-09-423-351C-5  
Perfect score: 81  
Sequence: 1 PRHQGVVMGMGQKDS 15  
Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 283416 seqs, 96216763 residues  
Total number of hits satisfying chosen parameters: 283416

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000  
Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : PIR\_79:.\*  
1: pir1.\*  
2: pir2.\*  
3: pir3.\*  
4: pir4.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	81	100.0	86	C43616	actin beta, cytosol
2	81	100.0	137	A28258	actin 5C, fruit f
3	81	100.0	158	D149465	alpha-cardiac acti
4	81	100.0	213	A61043	actin CA15 - sea s
5	81	100.0	308	A03000	actin 3 - fruit fl
6	81	100.0	328	S05430	actin beta - grass
7	81	100.0	349	B25819	actin, fetal skele
8	81	100.0	362	A26559	actin type 5, cyto
9	81	100.0	374	ATBOB	actin beta - bovin
10	81	100.0	374	1 ATBOG	actin gamma - bovi
11	81	100.0	374	2 JC5818	gamma-actin - huma
12	81	100.0	375	1 ATBOSM	actin, aortic smoo
13	81	100.0	375	1 ATRB	actin, skeletal mu
14	81	100.0	375	1 ATKTC	actin beta - rat
15	81	100.0	375	1 A48324	actin beta, cytosk
16	81	100.0	375	1 ATCHB	actin beta - chick
17	81	100.0	375	1 ATHUB	actin beta - human
18	81	100.0	375	1 ATHUG	actin gamma 1 - hu
19	81	100.0	375	1 ATMSB	actin beta - mouse
20	81	100.0	375	1 ATMSG	actin gamma - mous
21	81	100.0	375	1 ATRB	actin beta, non-mu
22	81	100.0	375	1 S11222	actin gamma, cytos
23	81	100.0	375	2 T25272	hypothetical prote
24	81	100.0	375	2 S71125	actin beta-2, cyto
25	81	100.0	375	2 S71124	actin beta-1, cyto
26	81	100.0	375	2 S71126	actin beta, cytosol
27	81	100.0	375	2 A55001	actin beta - goose
28	81	100.0	375	2 A54728	actin alpha, cardi
29	81	100.0	375	2 S42103	actin - Puccinia g

actin - Phaffia rh  
actin gamma, cytos  
actin gamma, smoot  
actin 8 - fruit fl  
actin CyI - sea ur  
Actin-1A - nematod  
actin 87E - fruit  
actin - fruit fly  
actin, cytosolic -  
actin, muscle - st  
actin 15A - sea ur  
actin - sea urchin  
actin gamma, enter  
actin - Hydra atte  
actin (clone gen3)  
actin (clones Ia a

## ALIGNMENTS

## RESULT 1

C43616  
actin beta, cytosolic - chicken (fragment)  
C:Species: Gallus gallus (chicken)  
C:Date: 11-Dec-1992 #sequence\_revision 11-Dec-1992 #text\_change 09-Jul-2004  
C:Accession: C43616  
R:Paterson, B.M.; Eldridge, J.D.  
Science 224, 1436-1438, 1984  
A:Title: alpha-Cardiac actin is the major sarcomeric isoform expressed in embryonic avia  
A:Reference number: A43616; MUID:84223949; PMID:6729461  
A:Accession: C43616  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-86 <PAT>  
A:Cross-references: UNIPROT:Q90736; GB:K02173; NID:G211054; PIDN:AAA98513.1; PID:G211055  
C:Superfamily: actin  
C:Keywords: cytosol; methylated amino acid  
F:73/Modified site: 3'-methylhistidine (His) #status predicted

Query Match 100.0%; Score 81; DB 2; Length 86;  
Best Local Similarity 100.0%; Pred. No. 3.3e-07;  
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 PRHQGVVMGMGQKDS 15  
Db 38 PRHQGVVMGMGQKDS 52

## RESULT 2

A28258  
actin 5C - fruit fly (Drosophila melanogaster)  
C:Species: Drosophila melanogaster  
C:Date: 30-Sep-1989 #sequence\_revision 30-Sep-1989 #text\_change 09-Jul-2004  
C:Accession: A28258  
R:Vigoreaux, J.O.; Tobin, S.L.  
Genes Dev. 1, 1161-1171, 1987  
A:Title: Stage-specific selection of alternative transcriptional initiation sites from t  
A:Reference number: A28258; MUID:88112795; PMID:3123314  
A:Accession: A28258  
A:Molecule type: mRNA  
A:Residues: 1-137 <VIG>  
A:Cross-references: UNIPROT:P10987  
A:Note: the authors translated the codon GAG for residue 96 as Gly  
C:Genetics:  
A:Gene: FlyBase:Act5C  
A:Cross-references: FlyBase:FBgn0000042  
C:Superfamily: actin  
C:Keywords: methylated amino acid  
F:74/Modified site: 3'-methylhistidine (His) #status predicted

Query Match 100.0%; Score 81; DB 2; Length 137;  
Best Local Similarity 100.0%; Pred. No. 5.3e-07;

Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 PRHQGVVVGMGQKDS 15  
| | | | | | | | | | | | | | | |  
Db 39 PRHQGVVVGMGQKDS 53

## RESULT 3

149465  
A;Title: alpha-cardiac actin - mouse (fragment)  
C;Species: Mus musculus (house mouse)  
C;Date: 02-Jul-1996 #sequence\_revision 02-Jul-1996 #text\_change 09-Jul-2004  
C;Accession: I49465; MUID:87053822; PMID:3023046  
R;Garner, I.; Minty, A.J.; Alonso, S.; Barton, P.J.R.; Buckingham, M.E.  
EMBO J. 5, 2559-2567, 1986  
A;Title: A 5' duplication of the alpha-cardiac actin gene in BALB/c mice is associated with a 5' deletion in the beta-cardiac actin gene  
A;Reference number: I49465; MUID:87053822; PMID:3023046  
A;Accession: I49465  
A;Status: translated from GB/EMBL/DBJ  
A;Molecule type: DNA  
A;Residues: 1-158 <RES>  
A;Cross-references: UNIPROT:Q61273; GB:M26775; NID:g191646; PIDN:AAA37165.1; PID:g553858  
A;Accession: I49466  
A;Status: translated from GB/EMBL/DBJ  
A;Molecule type: DNA  
A;Residues: 8-158 <RES>  
A;Cross-references: GB:M26776; NID:g191649; PIDN:AAA37166.1; PID:g553859  
A;Experimental source: adult cardiac muscle, BALB/c mice  
C;Genetics:  
A;Introns: 50/3  
C;Superfamily: actin  
C;Keywords: cardiac muscle; heart

Query Match 100.0%; Score 81; DB 2; Length 158;  
Best Local Similarity 100.0%; Pred. No. 6.2e-07;  
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 PRHQGVVVGMGQKDS 15  
| | | | | | | | | | | | | | | |  
Db 47 PRHQGVVVGMGQKDS 61

## RESULT 4

A61043  
actin CAL5 - sea squirt (Styela clava) (fragments)  
C;Species: Styela clava  
C;Date: 31-Dec-1993 #sequence\_revision 31-Dec-1993 #text\_change 09-Jul-2004  
C;Accession: A61043  
R;Beach, R.L.; Jeffery, W.R.  
Dev. Genet. 11, 2-14, 1990  
A;Title: Temporal and spatial expression of a cytoskeletal actin gene in the ascidian Styela clava  
A;Reference number: A61043; MUID:90298580; PMID:2361333  
A;Accession: A61043  
A;Status: not compared with conceptual translation  
A;Molecule type: mRNA  
A;Residues: 1-213 <BEA>  
A;Cross-references: UNIPROT:Q7M3V7  
C;Comment: This sequence is expressed in cells undergoing rapid cell division.  
C;Superfamily: actin  
C;Keywords: cytoskeleton; methylated amino acid; mitosis; structural protein  
F;73/Modified site: 3'-methylhistidine (His) #status predicted

Query Match 100.0%; Score 81; DB 2; Length 213;  
Best Local Similarity 100.0%; Pred. No. 8.4e-07;  
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 PRHQGVVVGMGQKDS 15  
| | | | | | | | | | | | | | | |  
Db 38 PRHQGVVVGMGQKDS 52

## RESULT 5

A03000

actin 3 - fruit fly (Drosophila melanogaster) (fragments)  
C;Species: Drosophila melanogaster  
C;Date: 31-Mar-1991 #sequence\_revision 31-Mar-1991 #text\_change 09-Jul-2004  
C;Accession: A03000  
R;Fyrberg, E.A.; Bond, B.J.; Hershey, N.D.; Mixter, K.S.; Davidson, N.  
Cell 24, 107-116, 1981  
A;Title: The actin genes of Drosophila: protein coding regions are highly conserved but ,  
A;Reference number: A03000; MUID:81210174; PMID:6263481  
A;Accession: A03000  
A;Molecule type: DNA  
A;Residues: 1-308 <FYR>  
A;Cross-references: UNIPROT:P02572  
A;Note: there are 68 unsequenced residues between positions 160 and 161. Partial sequence  
A;Note: the authors translated the codon GTT for residue 263 as Ile  
C;Genetics:  
A;Gene: FlyBase:Act42A  
A;Cross-references: FlyBase:FBgn0000043  
A;Map position: 42A  
C;Superfamily: actin  
C;Keywords: methylated amino acid  
F;74/Modified site: 3'-methylhistidine (His) #status predicted

Query Match 100.0%; Score 81; DB 2; Length 308;  
Best Local Similarity 100.0%; Pred. No. 1.2e-06;  
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 PRHQGVVVGMGQKDS 15  
| | | | | | | | | | | | | | | |  
Db 39 PRHQGVVVGMGQKDS 53

## RESULT 6

S05430  
actin beta - grass carp  
C;Species: Ctenopharyngodon idella (grass carp)  
C;Date: 07-Jun-1990 #sequence\_revision 07-Jun-1990 #text\_change 09-Jul-2004  
C;Accession: S05430  
R;Liu, Z.; Zhu, Z.; Roberg, K.; Faras, A.J.; Guise, K.S.; Kapuscinski, A.R.; Hackett, P.E.  
Nucleic Acids Res. 17, 5850, 1989  
A;Title: The beta-actin gene of carp (Ctenopharyngodon idella).  
A;Reference number: S05430; MUID:89345185; PMID:2762162  
A;Accession: S05430  
A;Status: translation not shown  
A;Molecule type: DNA  
A;Residues: 1-328 <LIU>  
A;Cross-references: UNIPROT:P83751; EMBL:M25013  
C;Genetics:  
A;Introns: 41/3; 121/3; 268/1  
C;Superfamily: actin  
C;Keywords: cytoskeleton; methylated amino acid  
F;73/Modified site: 3'-methylhistidine (His) #status predicted

Query Match 100.0%; Score 81; DB 2; Length 328;  
Best Local Similarity 100.0%; Pred. No. 1.3e-06;  
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 PRHQGVVVGMGQKDS 15  
| | | | | | | | | | | | | | | |  
Db 38 PRHQGVVVGMGQKDS 52

## RESULT 7

B25819  
actin, fetal skeletal/adult cardiac muscle - mouse (fragment)  
C;Species: Mus musculus (house mouse)  
C;Date: 23-Aug-1987 #sequence\_revision 23-Aug-1987 #text\_change 09-Jul-2004  
C;Accession: B25819  
R;Alonso, S.; Minty, A.; Bourlet, Y.; Buckingham, M.  
J. Mol. Evol. 23, 11-22, 1986  
A;Title: Comparison of three actin-coding sequences in the mouse; evolutionary relations  
A;Reference number: A25819; MUID:86200234; PMID:3084797  
A;Accession: B25819  
A;Molecule type: mRNA



A;Residues: 1-349 <ALO>  
A;Cross-references: UNIPROT:Q61275; GB:X03767; GB:J00381; GB:M10652; NID:G49869; PIDN:CA  
C;Superfamily: actin  
C;Keywords: cardiac muscle; heart; methylated amino acid; muscle; skeletal muscle  
F;7/Modified site: 3'-methylhistidine (His) #status predicted

Query Match 100.0%; Score 81; DB 2; Length 349;  
Best Local Similarity 100.0%; Pred. No. 1.4e-06;  
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 PRHQGVVMVGQKDS 15  
|||||  
DB 12 PRHQGVVMVGQKDS 26  
|||||

RESULT 8  
A26559  
actin type 5, cytosolic - chicken  
C;Species: Gallus gallus (chicken)  
C;Date: 09-Sep-1987 #sequence\_revision 09-Sep-1987 #text\_change 05-Dec-1997  
C;Accession: A26559  
R;Bergsma, D.J.; Chang, K.S.; Schwartz, R.J.  
Mol. Cell. Biol. 5, 1151-1162, 1985  
A;Reference number: A26559; MUID:85213487; PMID:4000121  
A;Accession: A26559  
A;Molecule type: DNA  
A;Residues: 1-362 <BER>  
C;Superfamily: actin  
C;Keywords: cytosol; methylated amino acid  
F;7/Modified site: 3'-methylhistidine (His) #status predicted

Query Match 100.0%; Score 81; DB 2; Length 362;  
Best Local Similarity 100.0%; Pred. No. 1.5e-06;  
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 PRHQGVVMVGQKDS 15  
|||||  
DB 39 PRHQGVVMVGQKDS 53  
|||||

RESULT 9  
ATBOB  
actin beta - bovine (tentative sequence)  
C;Species: Bos primigenius taurus (cattle)  
C;Date: 30-Sep-1991 #sequence\_revision 30-Sep-1991 #text\_change 09-Jul-2004  
C;Accession: E4185; A39105; A02999; A14185  
R;Vandekerckhove, J.; Weber, K.  
Eur. J. Biochem. 90, 451-462, 1978  
A;Title: Actin amino-acid sequences. Comparison of actins from calf thymus, bovine brain  
A;Reference number: A14185; MUID:79045349; PMID:213279  
A;Accession: E4185  
A;Molecule type: protein  
A;Residues: 1-374 <VAN>  
A;Cross-references: UNIPROT:P60712  
A;Note: only peptides that differed in composition from the corresponding peptides of ra  
R;Degen, J.L.; Neubauer, M.G.; Degen, S.J.F.; Seyfried, C.E.; Morris, D.R.  
J. Biol. Chem. 258, 12153-12162, 1983  
A;Title: Regulation of protein synthesis in mitogen-activated bovine lymphocytes. Analys  
A;Reference number: A39105; MUID:84032385; PMID:6195151  
A;Accession: A39105  
A;Molecule type: mRNA  
A;Residues: 76-227;344-374 <DEG>  
A;Cross-references: GB:K00622; GB:K00623  
A;Note: actins beta and gamma were not distinguished in this study  
C;Comment: Vertebrate nonmuscle cells contain two highly conserved cytoskeletal actins,  
C;Superfamily: actin  
C;Keywords: blocked amino end; cell motility; cytoskeleton; methylated amino acid; micro  
F;1/Modified site: blocked amino end (Asp) (probably acetylated) #status experimental  
F;7/Modified site: 3'-methylhistidine (His) #status predicted

Query Match 100.0%; Score 81; DB 1; Length 374;  
Best Local Similarity 100.0%; Pred. No. 1.5e-06;  
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 PRHQGVVMVGQKDS 15  
|||||  
DB 37 PRHQGVVMVGQKDS 51  
|||||

RESULT 10  
ATBOG  
actin gamma - bovine (tentative sequence)  
C;Species: Bos primigenius taurus (cattle)  
C;Date: 30-Sep-1991 #sequence\_revision 30-Sep-1991 #text\_change 09-Jul-2004  
C;Accession: E4185; A02999  
R;Vandekerckhove, J.; Weber, K.  
Eur. J. Biochem. 90, 451-462, 1978  
A;Title: Actin amino-acid sequences. Comparison of actins from calf thymus, bovine brain  
A;Reference number: A14185; MUID:79045349; PMID:213279  
A;Accession: E4185  
A;Molecule type: protein  
A;Residues: 1-374 <VAN>  
A;Cross-references: UNIPROT:P02571  
A;Note: only peptides that differed in composition from the corresponding peptides of ra  
C;Comment: Vertebrate nonmuscle cells contain two highly conserved cytoskeletal actins,  
C;Superfamily: actin  
C;Keywords: blocked amino end; cell motility; cytoskeleton; methylated amino acid; micro  
F;1/Modified site: blocked amino end (Glu) (probably acetylated) #status experimental  
F;7/Modified site: 3'-methylhistidine (His) #status predicted

Query Match 100.0%; Score 81; DB 1; Length 374;  
Best Local Similarity 100.0%; Pred. No. 1.5e-06;  
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 PRHQGVVMVGQKDS 15  
|||||  
DB 37 PRHQGVVMVGQKDS 51  
|||||

RESULT 11  
JCS818  
gamma-actin - human  
C;Species: Homo sapiens (man)  
C;Date: 20-Feb-1998 #sequence\_revision 13-Mar-1998 #text\_change 09-Jul-2004  
C;Accession: JC5818; PC4501  
R;Hauschildt, S.; Schwarz, C.; Heine, H.; Ulmer, A.J.; Flad, H.D.; Rietschel, E.T.; Jens  
Biochem. Biophys. Res. Commun. 241, 670-674, 1997  
A;Title: Actin: A target of lipopolysaccharide-induced phosphorylation in human monocytes  
A;Reference number: JC5818; MUID:98096379; PMID:9434766  
A;Accession: JC5818  
A;Molecule type: protein  
A;Residues: 1-374 <HAU>  
A;Cross-references: UNIPROT:P02571  
A;Experimental source: monocyte  
A;Accession: PC4501  
A;Molecule type: protein  
A;Residues: 1-61;84-112;147-190;196-209;215-253;335-358 <HA2>  
A;Experimental source: monocyte  
C;Comment: This protein is involved in a signal transduction that eventually leads to mo  
C;Superfamily: actin

Query Match 100.0%; Score 81; DB 2; Length 374;  
Best Local Similarity 100.0%; Pred. No. 1.5e-06;  
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 PRHQGVVMVGQKDS 15  
|||||  
DB 37 PRHQGVVMVGQKDS 51  
|||||

RESULT 12  
ATBOB  
actin, aortic smooth muscle - bovine  
C;Species: Bos primigenius taurus (cattle)  
C;Date: 30-Jun-1987 #sequence\_revision 30-Jun-1987 #text\_change 06-Sep-1996  
C;Accession: A02997; S13480

R;Vandekerckhove, J.; Weber, K.  
Differentiation 14, 123-133, 1979  
A;Title: The complete amino acid sequence of actins from bovine aorta, bovine heart, bovine skeletal muscle.  
A;Reference number: A02997; MUID:80047657; PMID:499690  
A;Accession: A02997  
A;Molecule type: protein  
A;Residues: 1-375 <VAN>  
R;Zevgolis, V.G.; Sotiroudis, T.G.; Evangelopoulos, A.E.  
Biochim. Biophys. Acta 1091, 222-230, 1991  
A;Title: Phosphorylase kinase from bovine stomach smooth muscle: a Ca(2+)-dependent protein.  
A;Reference number: S13480; MUID:91137633; PMID:1995080  
A;Accession: S13480  
A;Molecule type: protein  
A;Residues: 40-49 <ZEV>  
A;Experimental source: stomach  
A;Note: this material appears to be actin of aortic smooth muscle type or a related molecule.  
C;Superfamily: actin  
C;Keywords: acetylated amino end; methylated amino acid; muscle contraction  
F;1/Modified site: acetylated amino end (Glu) #status predicted  
F;73/Modified site: 3'-methylhistidine (His) #status experimental

Query Match 100.0%; Score 81; DB 1; Length 375;  
Best Local Similarity 100.0%; Pred. No. 1.5e-06;  
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 PRHQGVVMVGMGQKDS 15  
Db 38 PRHQGVVMVGMGQKDS 52

RESULT 13

actin, skeletal muscle - rabbit  
N;Alternate names: F-actin  
C;Species: Oryctolagus cuniculus (domestic rabbit)  
C;Date: 30-Sep-1991 #sequence revision 30-Sep-1991 #text change 07-May-1999  
A;Accession: A92182; A14185; S65873; S70610; S44393; A02994  
R;Collins, J.H.; Elzinga, M.  
J. Biol. Chem. 250, 5915-5920, 1975  
A;Title: The primary structure of actin from rabbit skeletal muscle. Completion and analysis of the primary structure.  
A;Reference number: A92182; MUID:75211334; PMID:1150665  
A;Accession: A92182  
A;Molecule type: protein  
A;Residues: 1-2, 'T', 'D', '6-11', 'D', '13-73', 'W', '74-78, 80-234, 236-308, 'T', 310-375 <COL>  
A;Note: this is the final paper in a series  
A;Note: this sequence has been revised in references A14185 and A90406  
R;Vandekerckhove, J.; Weber, K.  
Eur. J. Biochem. 90, 451-462, 1978  
A;Title: Actin amino-acid sequences. Comparison of actins from calf thymus, bovine brain and rabbit skeletal muscle.  
A;Reference number: A14185; MUID:79045349; PMID:213279  
A;Accession: A14185  
A;Molecule type: protein  
A;Residues: 1-18; 69-84 <VAN>  
R;Lu, R.C.; Elzinga, M.  
Biochemistry 15, 5801-5806, 1977  
A;Title: Partial amino acid sequence of brain actin and its homology with muscle actin.  
A;Reference number: A90406; MUID:78060866; PMID:588555  
A;Contents: annotation  
A;Note: residue 235 has been added and residue 309 has been revised  
R;Vandat, A.; Miller, C.; Phillips, M.; Muhllrad, A.; Reisler, E.  
FEBS Lett. 365, 149-151, 1995  
A;Title: A novel 27/16 kDa form of subunitin. cleaved actin: structural and functional characterization.  
A;Reference number: S65873; MUID:95300963; PMID:7781768  
A;Accession: S65873  
A;Molecule type: protein  
A;Residues: 235-241 <VAN>  
R;Strzelecka-Golaszewska, H.; Wozniak, A.; Hult, T.; Lindberg, U.  
Biochem. J. 316, 713-721, 1996  
A;Title: Effects of the type of divalent cation, Ca2+ or Mg2+, bound at the high-affinity site on the actin monomer.  
A;Reference number: S70610; MUID:96265033; PMID:8670143  
A;Accession: S70610  
A;Molecule type: protein

A;Residues: 48-54; 68-72; 235-243 <STR>  
A;Experimental source: skeletal muscle  
R;Bertrand, R.; Derancourt, J.; Kassab, R.  
FEBS Lett. 345, 113-119, 1994  
A;Title: The covalent maleimidebenzoyl-actin-myosin head complex. Cross-linking of the 5' and 3' ends of the actin monomer.  
A;Reference number: S44393; MUID:94259162; PMID:8200441  
A;Accession: S44393  
A;Molecule type: protein  
A;Residues: 48-64 <BER>  
A;Experimental source: skeletal muscle  
C;Superfamily: actin  
C;Keywords: acetylated amino end; ATP binding; methylated amino acid; muscle contraction  
F;1-375/Product: actin #status experimental <MAT>  
F;1/Modified site: acetylated amino end (Asp) #status experimental  
F;73/Modified site: 3'-methylhistidine (His) #status experimental

Query Match 100.0%; Score 81; DB 1; Length 375;  
Best Local Similarity 100.0%; Pred. No. 1.5e-06;  
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 PRHQGVVMVGMGQKDS 15  
Db 38 PRHQGVVMVGMGQKDS 52

RESULT 14

actin beta - rat  
C;Species: Rattus norvegicus (Norway rat)  
C;Date: 13-Jun-1983 #sequence revision 13-Jun-1983 #text change 22-Jun-1999  
A;Accession: A38571; A02999  
R;Nudel, U.; Zakut, R.; Shani, M.; Neuman, S.; Levy, Z.; Yaffe, D.  
Nucleic Acids Res. 11, 1759-1771, 1983  
A;Title: The nucleotide sequence of the rat cytoplasmic beta-actin gene.  
A;Reference number: A38571; MUID:83168920; PMID:6300777  
A;Accession: A38571  
A;Molecule type: DNA  
A;Residues: 1-375 <NDU>  
A;Cross-references: GB:J00691; NID:g202653; PIDN:AAA40657.1; PID:g202654  
C;Comment: Vertebrate nonmuscle cells contain two highly conserved cytoskeletal actins,  $\alpha$ -actin and  $\beta$ -actin.  
C;Genetics: 41/3; 121/3; 268/1; 328/3  
C;Introns: 41/3; 121/3; 268/1; 328/3  
C;Superfamily: actin  
C;Keywords: cell motility; cytoskeleton; methylated amino acid; microfilament; mitosis; cell division  
F;2-375/Product: actin beta #status predicted <MAT>  
F;73/Modified site: 3'-methylhistidine (His) #status predicted

Query Match 100.0%; Score 81; DB 1; Length 375;  
Best Local Similarity 100.0%; Pred. No. 1.5e-06;  
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 PRHQGVVMVGMGQKDS 15  
Db 38 PRHQGVVMVGMGQKDS 52

RESULT 15

actin beta, cytoskeletal - common carp  
C;Species: Cyprinus carpio (common carp)  
C;Date: 03-Feb-1994 #sequence revision 11-Apr-1997 #text change 09-Jul-2004  
A;Accession: A48324  
R;Liu, Z.; Zhu, Z.; Roberg, K.; Faras, A.; Guise, K.; Kapuscinski, A.R.; Hackett, P.B.; DNA Seq. 1, 125-136, 1990  
A;Title: Isolation and characterization of beta-actin gene of carp (Cyprinus carpio).  
A;Reference number: A48324; MUID:92190540; PMID:2134183  
A;Accession: A48324  
A;Molecule type: DNA  
A;Residues: 1-375 <LIU>  
A;Cross-references: UNIPROT:P83750; GB:M24113; NID:g213041; PIDN:AAA68886.1; PID:g213042  
A;Note: the authors translated the codon TTC for residue 21 as Pro, AAG for residue 50 as Pro  
A;Note: the authors failed to translated the codon GGT for residue 42 as Gly

C;Comment: Vertebrate nonmuscle cells contain two highly conserved cytoskeletal actins,  
 C;Genetics:  
 A;Introns: 41/3; 121/3; 268/3; 328/3  
 C;Superfamily: actin  
 C;Keywords: acetylated amino end; cell motility; cytoskeleton; methylated amino acid; mi  
 F;2-375/Product: actin beta, cytoskeletal #status predicted <MAT>  
 F;2/Modified site: acetylated amino end (Asp) (in mature form) #status predicted  
 F;73/Modified site: 3'-methylhistidine (His) #status predicted

Query Match 100.0%; Score 81; DB 1; Length 375;  
 Best Local Similarity 100.0%; Pred. No. 1.5e-06;  
 Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 PRHQGVVVGQKDS 15  
 |||||  
 Db 38 PRHQGVVVGQKDS 52

Search completed: April 8, 2005, 10:53:05  
 Job time : 9.3333 secs

**THIS PAGE BLANK (USPTO)**

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: April 8, 2005, 09:58:56 ; Search time 57.8667 Seconds  
(without alignments)  
132.739 Million cell updates/sec

Title: US-09-423-351C-5  
Perfect score: 81  
Sequence: 1 PRHQGVVMGQKDS 15

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 1612378 seqs, 512079187 residues

Total number of hits satisfying chosen parameters: 1612378

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : UniProt\_03:\*

1: uniprot\_sprot:\*

2: uniprot\_trembl:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	81	100.0	26	2 Q6Q298	Q6Q298 ovix aries
2	81	100.0	53	2 Q6AW41	Q6AW41 bombyx mori
3	81	100.0	78	2 O61375	O61375 schmidtea p
4	81	100.0	78	2 O61376	O61376 schmidtea p
5	81	100.0	78	2 O61377	O61377 schmidtea p
6	81	100.0	79	2 Q95L83	Q95L83 oryctolagus
7	81	100.0	80	2 Q9TTW4	Q9TTW4 bos taurus
8	81	100.0	86	2 Q90736	Q90736 gallus gall
9	81	100.0	91	2 O7ZZI1	O7ZZI1 salvelinus
10	81	100.0	96	2 Q86CV4	Q86CV4 drosophila
11	81	100.0	96	2 Q86FV7	Q86FV7 drosophila
12	81	100.0	99	2 Q9QZB4	Q9QZB4 cavia porce
13	81	100.0	100	2 Q862L0	Q862L0 bos taurus
14	81	100.0	104	2 Q86SB7	Q86SB7 gryllus bim
15	81	100.0	108	2 O18548	O18548 lytechinus
16	81	100.0	108	2 O18550	O18550 lytechinus
17	81	100.0	107	2 Q8T6P7	Q8T6P7 elysia chlo
18	81	100.0	121	2 Q8JG71	Q8JG71 oreochromis
19	81	100.0	121	2 Q78CQ1	Q78CQ1 pseudopleur
20	81	100.0	129	2 Q6VBE4	Q6VBE4 locusta mig
21	81	100.0	130	2 Q93298	Q93298 pseudopleur
22	81	100.0	139	2 Q9DFK2	Q9DFK2 gallicthys
23	81	100.0	144	2 Q90YX9	Q90YX9 fundulus he
24	81	100.0	150	2 Q804Y9	Q804Y9 ictalurus p
25	81	100.0	151	2 Q61274	Q61274 mus musculu
26	81	100.0	151	2 Q9PSV5	Q9PSV5 oryzias lat
27	81	100.0	158	2 Q61273	Q61273 mus musculu
28	81	100.0	164	1 ACT_SPOLI	Q11212 spodoptera
29	81	100.0	165	2 O17477	O17477 haliotis ru
30	81	100.0	181	2 Q8JG41	Q8JG41 dicentrarch
31	81	100.0	181	2 Q8AWC8	Q8AWC8 chionodraco

32 81 100.0 181 2 Q8AWZ3 Q8AWZ3 argyrolelec  
33 81 100.0 181 2 Q9DEU3 Q9DEU3 scophthalmu  
34 81 100.0 184 2 Q11209 Q11209 canis fami  
35 81 100.0 186 2 Q8X190 Q8X190 paxillus in  
36 81 100.0 189 2 Q8SPX4 Q8SPX4 canis fami  
37 81 100.0 191 2 Q64G12 Q64G12 oxyuranus s  
38 81 100.0 198 2 Q6Y250 Q6Y250 pagrus majo  
39 81 100.0 202 2 Q810R5 Q810R5 mus musculu  
40 81 100.0 212 2 Q6ZYL2 Q6ZYL2 axion lusit  
41 81 100.0 213 2 Q7M3V7 Q7M3V7 styela clav  
42 81 100.0 225 2 Q64G13 Q64G13 oxyuranus s  
43 81 100.0 236 2 Q6RXK3 Q6RXK3 penaeus van  
44 81 100.0 254 2 Q7Z7J6 Q7Z7J6 homo sapien  
45 81 100.0 275 2 Q64I75 Q64I75 reticuliter

#### ALIGNMENTS

##### RESULT 1

Q6Q298 PRELIMINARY; PRT; 26 AA.  
AC Q6Q298;  
DT 05-JUL-2004 (TrEMBLrel. 27, Created)  
DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)  
DE Beta actin (Fragment).  
OS Ovis aries (Sheep).  
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;  
OC Caprinae; Ovis.  
OX NCBI\_TaxID=9940;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC TISSUE=Abomasum;  
RA Shanmuganathan T., Hickford J.G.H., Savill M., Sykes A.R.;  
RL Submitted (MAR-2004) to the EMBL/GenBank/DBJ databases.  
DR EMBL; AY566300; AAS68014.1; -;  
FT NON\_TER 1  
FT NON\_TER 26  
SQ SEQUENCE 26 AA; 2802 MW; C75DA20C92E36C4B CRC64;

Query Match 100.0%; Score 81; DB 2; Length 26;  
Best Local Similarity 100.0%; Pred. No. 1.1e-06;  
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 PRHQGVVMGQKDS 15

Db 9 PRHQGVVMGQKDS 23

##### RESULT 2

Q6AW41 PRELIMINARY; PRT; 53 AA.  
AC Q6AW41;  
DT 25-OCT-2004 (TrEMBLrel. 28, Created)  
DT 25-OCT-2004 (TrEMBLrel. 28, Last sequence update)  
DT 25-OCT-2004 (TrEMBLrel. 28, Last annotation update)  
DE Cytoplasmic actin (fragment).  
OS Bombyx mori (Silk moth).  
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;  
OC Neoptera; Endopterygota; Lepidoptera; Glossata; Ditrysia; Bombycoidea;  
OC Bombycidae; Bombyx.  
OX NCBI\_TaxID=7091;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=Daizo; TISSUE=Silk gland;  
RA Shiomi K., Kajitani Z., Nakagaki M., Yamashita O.;  
RT "Baculovirus-mediated efficient gene transfer into the central nervous  
system of the silkworm, Bombyx mori.";  
RL Nihon Sanehigaku Zasshi 72:149-155(2003).  
RN [2]  
RP SEQUENCE FROM N.A.

```

RC STRAIN=Daizo; TISSUE=Silk gland;
RA Shiomi K.;
RL Submitted (AUG-2004) to the EMBL/GenBank/DBJ databases.
DR EMBL; AB186491; BAB35130.1; -.
DR InterPro; IPR004000; Actin_like.
DR Pfam; PF00022; Actin; 1.
FT NON TER 53
SQ SEQUENCE 53 AA; 5465 MW; 227AA8B4872EB86 CRC64;

Query Match 100.0%; Score 81; DB 2; Length 53;
Best Local Similarity 100.0%; Pred. No. 2.le-06;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 PRHQGVVGMGQKDS 15
Db 39 PRHQGVVGMGQKDS 53

RESULT 3
O61375 PRELIMINARY; PRT; 78 AA.
AC O61375;
DT 01-AUG-1998 (TrEMBLrel. 07, Created)
DT 01-AUG-1998 (TrEMBLrel. 07, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE Actin 1 (Fragment).
GN Name=DpActi;
OS Schmidtea polychroa.
OC Eukaryota; Metazoa; Platyhelminthes; Turbellaria; Seriata; Tricladida;
OC Paludicola; Dugesidae; Schmidtea.
OX NCBI_TaxID=50054;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=98186847; PubMed=9518500; DOI=10.1093/nar/26.8.2031;
RA Fagotti A., Gabbiani G., Pascolini R., Neuville P.;
RT "Multiple isoform recovery (MIR)-PCR: a simple method for the
isolation of related mRNA isoforms.";
RL Nucleic Acids Res. 26:2031-2033(1998).
CC -!- FUNCTION: Actins are highly conserved proteins that are involved
in various types of cell motility and are ubiquitously expressed
in all eukaryotic cells (By similarity).
CC -!- SIMILARITY: Belongs to the actin family.
DR EMBL; AF027162; AAC38982.1; -.
DR HSSP; P02577; INM1.
DR GO; GO:0005884; C:actin filament; IEA.
DR GO; GO:0003774; F:motor activity; IEA.
DR GO; GO:0005200; F:structural constituent of cytoskeleton; IEA.
DR InterPro; IPR004001; Actin_like.
DR Pfam; PF00022; Actin; 1.
DR PRINTS; PR00190; ACTIN.
DR PROSITE; PS00406; ACTINS_1; 1.
KW Structural protein.
FT NON TER 78
SQ SEQUENCE 78 AA; 8205 MW; 572BAB2698C24404 CRC64;

Query Match 100.0%; Score 81; DB 2; Length 78;
Best Local Similarity 100.0%; Pred. No. 3e-06;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 PRHQGVVGMGQKDS 15
Db 39 PRHQGVVGMGQKDS 53

RESULT 4
O61376 PRELIMINARY; PRT; 78 AA.
AC O61376;
DT 01-AUG-1998 (TrEMBLrel. 07, Created)
DT 01-AUG-1998 (TrEMBLrel. 07, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE Actin 2 (Fragment).

```

```

GN Name=DpAct2;
OS Schmidtea polychroa.
OC Eukaryota; Metazoa; Platyhelminthes; Turbellaria; Seriata; Tricladida;
OC Paludicola; Dugesidae; Schmidtea.
OX NCBI_TaxID=50054;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=98186847; PubMed=9518500; DOI=10.1093/nar/26.8.2031;
RA Fagotti A., Gabbiani G., Pascolini R., Neuville P.;
RT "Multiple isoform recovery (MIR)-PCR: a simple method for the
isolation of related mRNA isoforms.";
RL Nucleic Acids Res. 26:2031-2033(1998).
CC -!- FUNCTION: Actins are highly conserved proteins that are involved
in various types of cell motility and are ubiquitously expressed
in all eukaryotic cells (By similarity).
CC -!- SIMILARITY: Belongs to the actin family.
DR EMBL; AF027162; AAC38982.1; -.
DR HSSP; P02577; INM1.
DR GO; GO:0005884; C:actin filament; IEA.
DR GO; GO:0003774; F:motor activity; IEA.
DR GO; GO:0005200; F:structural constituent of cytoskeleton; IEA.
DR InterPro; IPR004001; Actin_like.
DR Pfam; PF00022; Actin; 1.
DR PRINTS; PR00190; ACTIN.
DR PROSITE; PS00406; ACTINS_1; 1.
KW Structural protein.
FT NON TER 78
SQ SEQUENCE 78 AA; 8285 MW; B9DFF9B28E4ADB0E CRC64;

Query Match 100.0%; Score 81; DB 2; Length 78;
Best Local Similarity 100.0%; Pred. No. 3e-06;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 PRHQGVVGMGQKDS 15
Db 39 PRHQGVVGMGQKDS 53

RESULT 5
O61377 PRELIMINARY; PRT; 78 AA.
AC O61377;
DT 01-AUG-1998 (TrEMBLrel. 07, Created)
DT 01-AUG-1998 (TrEMBLrel. 07, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE Actin 3 (Fragment).
GN Name=DpAct3;
OS Schmidtea polychroa.
OC Eukaryota; Metazoa; Platyhelminthes; Turbellaria; Seriata; Tricladida;
OC Paludicola; Dugesidae; Schmidtea.
OX NCBI_TaxID=50054;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=98186847; PubMed=9518500; DOI=10.1093/nar/26.8.2031;
RA Fagotti A., Gabbiani G., Pascolini R., Neuville P.;
RT "Multiple isoform recovery (MIR)-PCR: a simple method for the
isolation of related mRNA isoforms.";
RL Nucleic Acids Res. 26:2031-2033(1998).
CC -!- FUNCTION: Actins are highly conserved proteins that are involved
in various types of cell motility and are ubiquitously expressed
in all eukaryotic cells (By similarity).
CC -!- SIMILARITY: Belongs to the actin family.
DR EMBL; AF027163; AAC38983.1; -.
DR HSSP; P02577; INM1.
DR GO; GO:0005884; C:actin filament; IEA.
DR GO; GO:0003774; F:motor activity; IEA.
DR GO; GO:0005200; F:structural constituent of cytoskeleton; IEA.
DR InterPro; IPR004001; Actin.
DR Pfam; PF00022; Actin; 1.
DR PRINTS; PR00190; ACTIN.
DR PROSITE; PS00406; ACTINS_1; 1.

```

```
KW Structural protein.
FT NON TER 78
SQ SEQUENCE 78 AA; 8299 MW; B9DFP1108E4ADBOE CRC64;

Query Match 100.0%; Score 81; DB 2; Length 78;
Best Local Similarity 100.0%; Pred. No. 3e-06;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 PRHQGVVMVGQKDS 15
Db 39 PRHQGVVMVGQKDS 53

RESULT 6
Q95L83 PRELIMINARY; PRT; 79 AA.
AC Q95L83;
DT 01-DEC-2001 (TrEMBLrel. 19, Created)
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE Beta-actin (Fragment)
OS Oryctolagus cuniculus (Rabbit).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Lagomorpha; Leporidae; Oryctolagus.
OX NCBI_TaxID=9986;
RN [1]
RP SEQUENCE FROM N.A. PubMed=12388084;
RX MEDLINE=22306422;
RA Wang D.-A., Du H., Jaggar J.H., Brindley D.N., Tigyi G.J.,
RA Watsky M.A.;
RT "Injury-elicited differential transcriptional regulation of
RT phospholipid growth factor receptors in the cornea.";
RL Am. J. Physiol. Cell Physiol. 283:C1646-C1654(2002).
CC -!- FUNCTION: Actins are highly conserved proteins that are involved
CC in various types of cell motility and are ubiquitously expressed
CC in all eukaryotic cells (By similarity).
CC -!- SIMILARITY: Belongs to the actin family.
DR EMBL; AF404278; AAL01885.1; -.
DR HSSP; P02577; INMI.
DR GO; GO:0005884; C:actin filament; IEA.
DR GO; GO:0003774; F:motor activity; IEA.
DR GO; GO:0005200; F:structural constituent of cytoskeleton; IEA.
DR InterPro; IPR004001; Actin.
DR InterPro; IPR004000; Actin_like.
DR Pfam; PF00022; Actin; 1.
DR PRINTS; PR00190; ACTIN.
DR PROSITE; PS00406; ACTINS_1; 1.
KW Structural protein.
FT NON TER 79
SQ SEQUENCE 79 AA; 8420 MW; 05056E4E2AC092F8 CRC64;

Query Match 100.0%; Score 81; DB 2; Length 79;
Best Local Similarity 100.0%; Pred. No. 3.1e-06;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 PRHQGVVMVGQKDS 15
Db 40 PRHQGVVMVGQKDS 54

RESULT 7
Q9TTW4 PRELIMINARY; PRT; 80 AA.
ID Q9TTW4
AC Q9TTW4;
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE Beta actin (Fragment).
OS Bos taurus (Bovine).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
OC Bovinae; Bos.
```

```
OX NCBI_TaxID=9913;
RN [1]
RP SEQUENCE FROM N.A.
RA John S.J., Bilocheau-Goeseels S.;
RL Submitted (OCT-1999) to the EMBL/GenBank/DBJ databases.
CC -!- FUNCTION: Actins are highly conserved proteins that are involved
CC in various types of cell motility and are ubiquitously expressed
CC in all eukaryotic cells (By similarity).
CC -!- SIMILARITY: Belongs to the actin family.
DR EMBL; AF191490; AAF05984.1; -.
DR HSSP; P02577; INMI.
DR GO; GO:0005884; C:actin filament; IEA.
DR GO; GO:0003774; F:motor activity; IEA.
DR GO; GO:0005200; F:structural constituent of cytoskeleton; IEA.
DR InterPro; IPR004001; Actin.
DR InterPro; IPR004000; Actin_like.
DR Pfam; PF00022; Actin; 1.
DR PRINTS; PR00190; ACTIN.
DR PROSITE; PS00406; ACTINS_1; 1.
KW Structural protein.
FT NON TER 80
SQ SEQUENCE 80 AA; 9133 MW; 47354ABB7010668D CRC64;

Query Match 100.0%; Score 81; DB 2; Length 80;
Best Local Similarity 100.0%; Pred. No. 3.1e-06;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 PRHQGVVMVGQKDS 15
Db 4 PRHQGVVMVGQKDS 18

RESULT 8
Q90736 PRELIMINARY; PRT; 86 AA.
ID Q90736
AC Q90736;
DT 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE Beta-actin (Fragment).
OS Gallus gallus (Chicken).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Archosauria; Aves; Neognathae; Galliformes; Phasianidae;
OC Gallus.
OX NCBI_TaxID=9031;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=84223949; PubMed=6729461;
RA Paterson B.M., Eldridge J.D.;
RT "alpha-Cardiac actin is the major sarcomeric isoform expressed in
RT embryonic avian skeletal muscle.";
RL Science 224:1436-1438(1984).
CC -!- FUNCTION: Actins are highly conserved proteins that are involved
CC in various types of cell motility and are ubiquitously expressed
CC in all eukaryotic cells (By similarity).
CC -!- SIMILARITY: Belongs to the actin family.
DR EMBL; K02173; AAA98513.1; -.
DR PIR; C43616; C43616.
DR HSSP; P02577; INMI.
DR GO; GO:0005884; C:actin filament; IEA.
DR GO; GO:0003774; F:motor activity; IEA.
DR GO; GO:0005200; F:structural constituent of cytoskeleton; IEA.
DR Pfam; PF00022; Actin; 1.
DR PRINTS; PR00190; ACTIN.
DR PROSITE; PS00406; ACTINS_1; 1.
KW Structural protein.
FT NON TER 86
SQ SEQUENCE 86 AA; 9351 MW; A55285196A328B6E CRC64;

Query Match 100.0%; Score 81; DB 2; Length 86;
Best Local Similarity 100.0%; Pred. No. 3.3e-06;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY 1 PRHQGVVWVGQKDS 15
|||||
Db 38 PRHQGVVWVGQKDS 52

RESULT 9
Q7ZZI1 PRELIMINARY; PRT; 91 AA.
AC Q7ZZI1;
DT 01-JUN-2003 (Tremblrel. 24, Created)
DT 01-JUN-2003 (Tremblrel. 24, Last sequence update)
DT 01-MAR-2004 (Tremblrel. 26, Last annotation update)
DE Beta-actin (Fragment).
OS Salvelinus alpinus (Arctic char).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Euteleostei;
OC Protacanthopterygii; Salmoniformes; Salmonidae; Salvelinus.
OX NCBI_TaxID=8036;
RN [1]
RP SEQUENCE FROM N.A.
RA McGowan C., Davidson E.A., Davidson W.S.;
RL Submitted (MAR-2003) to the EMBL/GenBank/DBJ databases.
CC -!- FUNCTION: Actins are highly conserved proteins that are involved
CC in various types of cell motility and are ubiquitously expressed
CC in all eukaryotic cells (By similarity).
CC -!- SIMILARITY: Belongs to the actin family.
CC in all eukaryotic cells (By similarity).
CC -!- SIMILARITY: Belongs to the actin family.
DR EMBL; AY262761; AAP31127.1; -.
DR HSSP; P02577; INMI.
DR GO; GO:0005884; C:actin filament; IEA.
DR GO; GO:0003774; F:motor activity; IEA.
DR GO; GO:0005200; F:structural constituent of cytoskeleton; IEA.
DR InterPro; IPR004001; Actin.
DR Pfam; PF00022; Actin; 1.
DR PRINTS; PR00190; ACTIN.
DR PROSITE; PS00406; ACTINS_1; 1.
DR Structural protein.
FT NON_TER 1
FT NON_TER 91
FT NON_TER 91
SQ SEQUENCE 91 AA; 10063 MW; 7118DB6663CD895C CRC64;

Query Match 100.0%; Score 81; DB 2; Length 91;
Best Local Similarity 100.0%; Pred. No. 3.5e-06;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 PRHQGVVWVGQKDS 15
|||||
Db 36 PRHQGVVWVGQKDS 50

RESULT 10
Q86CV4 PRELIMINARY; PRT; 96 AA.
AC Q86CV4;
DT 01-JUN-2003 (Tremblrel. 24, Created)
DT 01-JUN-2003 (Tremblrel. 24, Last sequence update)
DT 01-MAR-2004 (Tremblrel. 26, Last annotation update)
DE Actin E2 (Fragment).
GN Name=ActE2;
OS Drosophila novamexicana.
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
OX NCBI_TaxID=47314;
RN [1]
RP SEQUENCE FROM N.A.
RA STRAIN=15010-1031.0;
RX MEDLINE=22480340; PubMed=12574518; DOI=10.1073/pnas.0336368100;
RA Wittkopp P.J., Williams B.L., Selegue J.E., Carroll S.B.;
RT "Drosophila pigmentation evolution: divergent genotypes underlying
RT convergent phenotypes.";
Proc. Natl. Acad. Sci. U.S.A. 100:1808-1813(2003).
CC -!- FUNCTION: Actins are highly conserved proteins that are involved
CC in various types of cell motility and are ubiquitously expressed
CC in all eukaryotic cells (By similarity).
CC -!- SIMILARITY: Belongs to the actin family.
DR EMBL; AY165516; AAP12711.1; -.
DR HSSP; P02568; 1LCU.
DR FlyBase; FBgn0066281; Dame\ActE2.
DR GO; GO:0005884; C:actin filament; IEA.
DR GO; GO:0003774; F:motor activity; IEA.
DR GO; GO:0005200; F:structural constituent of cytoskeleton; IEA.
DR InterPro; IPR004001; Actin.
DR Pfam; PF00022; Actin; 1.
DR PRINTS; PR00190; ACTIN.
DR PROSITE; PS00406; ACTINS_1; 1.
DR Structural protein.
FT NON_TER 96
FT NON_TER 96
SQ SEQUENCE 96 AA; 10610 MW; 1454B7D9DD3F3564 CRC64;

Query Match 100.0%; Score 81; DB 2; Length 96;
Best Local Similarity 100.0%; Pred. No. 3.7e-06;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 PRHQGVVWVGQKDS 15
|||||
Db 39 PRHQGVVWVGQKDS 53

RESULT 11
Q86FV7 PRELIMINARY; PRT; 96 AA.
AC Q86FV7;
DT 01-JUN-2003 (Tremblrel. 24, Created)
DT 01-JUN-2003 (Tremblrel. 24, Last sequence update)
DT 01-MAR-2004 (Tremblrel. 26, Last annotation update)
DE Actin E2 (Fragment).
GN Name=ActE2;
OS Drosophila americana (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
OX NCBI_TaxID=40366;
RN [1]
RP SEQUENCE FROM N.A.
RA STRAIN=15010-0951.0;
RX MEDLINE=22480340; PubMed=12574518; DOI=10.1073/pnas.0336368100;
RA Wittkopp P.J., Williams B.L., Selegue J.E., Carroll S.B.;
RT "Drosophila pigmentation evolution: divergent genotypes underlying
RT convergent phenotypes.";
Proc. Natl. Acad. Sci. U.S.A. 100:1808-1813(2003).
CC -!- FUNCTION: Actins are highly conserved proteins that are involved
CC in various types of cell motility and are ubiquitously expressed
CC in all eukaryotic cells (By similarity).
CC -!- SIMILARITY: Belongs to the actin family.
DR EMBL; AY165516; AAP12711.1; -.
DR HSSP; P02568; 1LCU.
DR FlyBase; FBgn0066281; Dame\ActE2.
DR GO; GO:0005884; C:actin filament; IEA.
DR GO; GO:0003774; F:motor activity; IEA.
DR GO; GO:0005200; F:structural constituent of cytoskeleton; IEA.
DR InterPro; IPR004001; Actin.
DR Pfam; PF00022; Actin; 1.
DR PRINTS; PR00190; ACTIN.
DR PROSITE; PS00406; ACTINS_1; 1.
DR Structural protein.
FT NON_TER 96
FT NON_TER 96
SQ SEQUENCE 96 AA; 10610 MW; 1454B7D9DD3F3564 CRC64;

Query Match 100.0%; Score 81; DB 2; Length 96;
Best Local Similarity 100.0%; Pred. No. 3.7e-06;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```



```

Qy 1 PRHQGVVMGQKDS 15
Db 39 PRHQGVVMGQKDS 53

RESULT 12
Q9QZB4 PRELIMINARY; PRT; 99 AA.
AC Q9QZB4;
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE Cytoplasmic actin (Fragment).
OS Cavia porcellus (Guinea pig).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Hystricognathi; Caviidae; Cavia.
OX NCBI_TaxID=10141;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Adrenal gland;
RA Buelow H.E., Bernhardt R.;
RL Submitted (OCT-1999) to the EMBL/GenBank/DBJ databases.
CC -!- FUNCTION: Actins are highly conserved proteins that are involved
CC in various types of cell motility and are ubiquitously expressed
CC in all eukaryotic cells (By similarity).
CC -!- SIMILARITY: Belongs to the actin family.
DR EMBL; AF191277; AAF13923.1; -.
DR HSSP; P02577; INMI.
DR GO; GO:0005884; C:actin filament; IEA.
DR GO; GO:0003774; F:motor activity; IEA.
DR GO; GO:0005200; F:structural constituent of cytoskeleton; IEA.
DR InterPro; IPR004001; Actin.
DR InterPro; IPR004000; Actin_like.
DR Pfam; PF00022; Actin; 1.
DR PRINTS; PR00190; ACTIN.
DR PROSITE; PS00406; ACTINS_1; 1.
DR Structural protein.
FT NON TER 99
SQ SEQUENCE 99 AA; 10973 MW; 0562B645E8DAF17B CRC64;

Query Match 100.0%; Score 81; DB 2; Length 99;
Best Local Similarity 100.0%; Pred. No. 3.8e-06;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 PRHQGVVMGQKDS 15
Db 38 PRHQGVVMGQKDS 52

RESULT 13
Q862L0 PRELIMINARY; PRT; 100 AA.
AC Q862L0;
DT 01-JUN-2003 (TrEMBLrel. 24, Created)
DT 01-JUN-2003 (TrEMBLrel. 24, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Similar to b-actin (Fragment).
OS Bos taurus (Bovine).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
OC Bovinae; Bos.
OX NCBI_TaxID=9913;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=22544902; PubMed=12658628; DOI=10.1002/mrd.10292;
RA Ishiwata H., Katsuma S., Kizaki K., Patel O.V., Nakano H.,
RA Takahashi T., Inai K., Hirasawa A., Shiojima S., Ikawa H., Suzuki Y.,
RA Tsujimoto G., Izaike Y., Todoroki J., Hashizume K.;
RT "Characterization of gene expression profiles in early bovine
RT pregnancy using a custom cDNA microarray.";
RL Mol. Reprod. Dev. 65:9-18(2003).
CC -!- FUNCTION: Actins are highly conserved proteins that are involved
CC in various types of cell motility and are ubiquitously expressed

```

```

CC in all eukaryotic cells (By similarity).
CC -!- SIMILARITY: Belongs to the actin family.
DR EMBL; AB098974; BAC56464.1; -.
DR HSSP; P02568; IP8Z.
DR GO; GO:0005884; C:actin filament; IEA.
DR GO; GO:0003774; F:motor activity; IEA.
DR GO; GO:0005200; F:structural constituent of cytoskeleton; IEA.
DR InterPro; IPR004001; Actin.
DR InterPro; IPR004000; Actin_like.
DR Pfam; PF00022; Actin; 1.
DR PRINTS; PR00190; ACTIN.
DR PROSITE; PS00406; ACTINS_1; 1.
DR PROSITE; PS01132; ACTINS_ACT_LIKE; 1.
KW Structural protein.
FT NON TER 1
FT NON TER 100
SQ SEQUENCE 100 AA; 11601 MW; 0E14674471D0019D CRC64;

Query Match 100.0%; Score 81; DB 2; Length 100;
Best Local Similarity 100.0%; Pred. No. 3.9e-06;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 PRHQGVVMGQKDS 15
Db 1 PRHQGVVMGQKDS 15

RESULT 14
Q86SB7 PRELIMINARY; PRT; 104 AA.
AC Q86SB7;
DT 01-JUN-2003 (TrEMBLrel. 24, Created)
DT 01-JUN-2003 (TrEMBLrel. 24, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Actin (Fragment).
DE Name=ACT;
OS Gryllus bimaculatus (Two-spotted cricket).
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC Neoptera; Orthopteroidea; Orthoptera; Ensifera; Gryllidae; Gryllinae;
OC Gryllus.
OX NCBI_TaxID=6999;
RN [1]
RP SEQUENCE FROM N.A.
RA Zhang H., Shimmyo Y., Hirose A., Mito T., Inoue Y., Ohuchi H.,
RA Loukeris T.G., Noji S.;
RT "Extrachromosomal transposition of the transposable element Minos
RT occurs in embryos of the cricket Gryllus bimaculatus.";
RL Dev. Growth Differ. 0:0-0(2002).
CC -!- FUNCTION: Actins are highly conserved proteins that are involved
CC in various types of cell motility and are ubiquitously expressed
CC in all eukaryotic cells (By similarity).
CC -!- SIMILARITY: Belongs to the actin family.
DR EMBL; AB087882; BAC55093.1; -.
DR HSSP; P02577; INMI.
DR GO; GO:0005884; C:actin filament; IEA.
DR GO; GO:0003774; F:motor activity; IEA.
DR GO; GO:0005200; F:structural constituent of cytoskeleton; IEA.
DR InterPro; IPR004001; Actin.
DR InterPro; IPR004000; Actin_like.
DR Pfam; PF00022; Actin; 1.
DR PRINTS; PR00190; ACTIN.
DR PROSITE; PS00406; ACTINS_1; 1.
KW Structural protein.
FT NON TER 104
SQ SEQUENCE 104 AA; 11511 MW; 332777A2B08669B0 CRC64;

Query Match 100.0%; Score 81; DB 2; Length 104;
Best Local Similarity 100.0%; Pred. No. 4e-06;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 PRHQGVVMGQKDS 15
Db 39 PRHQGVVMGQKDS 53

```

Search completed: April 8, 2005, 12:03:08  
Job time : 57.8667 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: April 8, 2005, 09:58:56 ; Search time 62.4 seconds  
(without alignments)  
92.971 Million cell updates/sec

Title: US-09-423-351C-6  
Perfect score: 74  
Sequence: 1 TFTPMPMVAIQAVL 15

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 2105692 seqs, 386760381 residues

Total number of hits satisfying chosen parameters: 2105692

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : A\_Geneseq\_16Dec04.\*  
1: geneseqp1980s.\*  
2: geneseqp1990s.\*  
3: geneseqp2000s.\*  
4: geneseqp2001s.\*  
5: geneseqp2002s.\*  
6: geneseqp2003as.\*  
7: geneseqp2003bs.\*  
8: geneseqp2004s.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	74	100.0	15	2 AAW92532	Beta-acti
2	74	100.0	97	4 AAm13687	Peptide #
3	74	100.0	97	4 ABB32618	Peptide #
4	74	100.0	97	4 AAm26087	Peptide #
5	74	100.0	97	4 ABB27467	Human pep
6	74	100.0	97	4 ABB18116	Protein #
7	74	100.0	97	4 AAm65825	Human bon
8	74	100.0	97	4 AAm53448	Human bra
9	74	100.0	97	4 ABB47471	Human liv
10	74	100.0	97	4 AAm01437	Peptide #
11	74	100.0	97	5 ABB35459	Human pep
12	74	100.0	137	4 AAm33076	Novel hum
13	74	100.0	150	4 ABB15101	Novel hum
14	74	100.0	166	4 ABB26869	Novel hum
15	74	100.0	168	5 ABB42912	Human ova
16	74	100.0	204	5 ABB43016	Human ova
17	74	100.0	239	7 ABB31206	Human dia
18	74	100.0	274	6 ABB70549	Human adi
19	74	100.0	276	3 AAG50947	Arabidops
20	74	100.0	277	3 AAG05115	Arabidops
21	74	100.0	294	7 ABB3068	Human apo
22	74	100.0	304	3 AAG50946	Arabidops
23	74	100.0	305	3 AAG05114	Arabidops
24	74	100.0	321	3 AAG50945	Arabidops
25	74	100.0	322	3 AAG05113	Arabidops

## ALIGNMENTS

RESULT 1  
AAW92532  
ID AAW92532 standard; peptide; 15 AA.

AC AAW92532;  
XX  
DT 26-APR-1999 (first entry)  
XX  
DE Beta-actin reference peptide substrate #6.

XX Peptide substrate; CCT; eukaryotic type II chaperonin complex; cyclin;  
OS binding agent; substrate-binding site; SBS; substrate folding; actin;  
FN tubulin; treatment; cancer; anticancer drug; viral infection; screening;  
PD reduced toxicity.

XX Synthetic.

XX WO9853322-A1.

XX 26-NOV-1998.

XX 22-MAY-1998; 98WO-GB001485.

XX 23-MAY-1997; 97GB-00010762.

XX (CANC-) INST CANCER RES ROYAL CANCER HOSPITAL.

XX Willison K, Hynes G, Liou AK;

XX WPI, 1999-070162/06.

XX Identifying specific binding agents for substrate binding site in CCT  
PT chaperonin complex - also new peptide binding agents and their mimetics,  
PT and peptides containing a specific CCT binding site, used for treating  
cancer.

XX Disclosure; Fig 10; 97pp; English.

XX This invention describes a method which uses the CCT (eukaryotic type II  
CC chaperonin) complex or part of it, for identifying a binding agent that  
CC can occupy a substrate-binding site (SBS) on the CCT complex. By binding  
CC to the CCT complex, the binding agents block an SBS so that biological  
CC activity of the CCT complex is affected, particularly its ability to fold  
CC substrates such as actin, tubulin and cyclin. The binding agents are  
CC useful for treatment of cancer, particularly when used in combination  
CC with an anticancer drug, or viral infections. Nucleic acid fragments are  
CC used to screen for agents, e.g. binding agents that modulate interaction  
CC between the CCT complex and a protein that is to be folded. The binding

Aag50873 Arabidops  
Aag05583 Arabidops  
Aag50944 Arabidops  
Aag54243 Arabidops  
Aag06484 Arabidops  
Adi63011 Human apo  
Aag06483 Arabidops  
Aag50872 Arabidops  
Aag05582 Arabidops  
Aag54242 Arabidops  
Aag50943 Arabidops  
Aab12985 Human bet  
Aap61532 Sequence  
Abb77395 Human act  
Aab64271 Angiogene  
Aam04830 Rat cytop  
Adm85212 Rat actin  
Ade61174 Rat Prote  
Adf30525 Rat angio  
Adi63062 Human apo

CC agents may target cells that are actively synthesising tubulin etc.  
CC (unlike known microtubule-stabilising agents that affect all cells), so  
CC should have reduced toxicity for normal cells. AA92527-W92541 are  
CC peptide substrates used in the method of the invention  
XX  
XX  
SQ Sequence 15 AA;

Query Match 100.0%; Score 74; DB 2; Length 15;  
Best Local Similarity 100.0%; Pred. NO. 3.2e-07; Indels 0; Gaps 0;  
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 TTNTPAMYVAIQAVL 15  
Db 1 TTNTPAMYVAIQAVL 15

RESULT 2  
AA13687  
ID AA13687 standard; protein; 97 AA.  
XX  
AC AA13687;  
XX  
DT 12-OCT-2001 (first entry)  
DE Peptide #121 encoded by probe for measuring cervical gene expression.  
XX Probe; human; microarray; gene expression; cervical epithelial cell;  
KW cervical cancer.  
XX Homo sapiens.  
XX WO200157278-A2.  
XX  
PD 09-AUG-2001.

XX 30-JAN-2001; 2001WO-US000670.  
XX 04-FEB-2000; 2000US-0180312P.  
XX 26-MAY-2000; 2000US-0207456P.  
XX 30-JUN-2000; 2000US-00608408.  
XX 03-AUG-2000; 2000US-00632366.  
XX 21-SEP-2000; 2000US-0234687P.  
XX 27-SEP-2000; 2000US-0236359P.  
XX 04-OCT-2000; 2000GB-00024263.  
XX (MOLE-) MOLECULAR DYNAMICS INC.  
XX Penn SG, Hanzel DK, Chen W, Rank DR;  
XX WPI; 2001-488901/53.  
XX Human genome-derived single exon nucleic acid probes useful for analyzing  
XX gene expression in human cervical epithelial cells.  
XX  
XX Claim 27; SEQ ID NO 18513; 487pp; English.

XX The present invention relates to human single exon nucleic acid probes  
XX (SENPs: see A110068-AA128459). The present sequence is a peptide encoded  
XX by one such probe. The SENPs are derived from human HeLa cells. The SENPs  
XX can be used to produce a single exon microarray, which can be used for  
XX measuring human gene expression in a sample derived from human cervical  
XX epithelial cells. By measuring gene expression, the probes are therefore  
XX useful in grading and/or staging of diseases of the cervix, notably  
XX cervical cancer. Note: The sequence data for this patent did not form  
XX part of the printed specification, but was obtained in electronic format  
XX directly from WIPO at ftp.wipo.int/pub/published\_pct\_sequences  
XX  
XX Sequence 97 AA;

Query Match 100.0%; Score 74; DB 4; Length 97;  
Best Local Similarity 100.0%; Pred. NO. 3e-06; Indels 0; Gaps 0;  
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 TTNTPAMYVAIQAVL 15  
Db 5 TTNTPAMYVAIQAVL 19

RESULT 3  
ABB32618  
ID ABB32618 standard; peptide; 97 AA.  
XX  
AC ABB32618;  
XX  
DT 04-FEB-2002 (first entry)  
DE Peptide #124 encoded by human foetal liver single exon probe.  
XX Human; foetal liver; gene expression; single exon nucleic acid probe.  
XX Homo sapiens.  
XX WO200157277-A2.  
XX  
PD 09-AUG-2001.

XX 30-JAN-2001; 2001WO-US000669.  
XX 04-FEB-2000; 2000US-0180312P.  
XX 26-MAY-2000; 2000US-0207456P.  
XX 30-JUN-2000; 2000US-00608408.  
XX 03-AUG-2000; 2000US-00632366.  
XX 21-SEP-2000; 2000US-0234687P.  
XX 27-SEP-2000; 2000US-0236359P.  
XX 04-OCT-2000; 2000GB-00024263.  
XX (MOLE-) MOLECULAR DYNAMICS INC.  
XX Penn SG, Hanzel DK, Chen W, Rank DR;  
XX WPI; 2001-483447/52.

XX Human genome-derived single exon nucleic acid probes useful for analyzing  
XX gene expression in human fetal liver.  
XX  
XX Claim 27; SEQ ID NO 25253; 639pp + Sequence Listing; English.

XX The invention relates to a single exon nucleic acid probe for measuring  
XX human gene expression in a sample derived from human foetal liver. The  
XX single exon nucleic acid probes may be used for predicting, measuring and  
XX displaying gene expression in samples derived from human fetal liver. The  
XX present sequence is a peptide encoded by a single exon nucleic acid probe  
XX of the invention. Note: The sequence data for this patent did not form  
XX part of the printed specification, but was obtained in electronic format  
XX directly from WIPO at ftp.wipo.int/pub/published\_pct\_sequences  
XX  
XX Sequence 97 AA;

Query Match 100.0%; Score 74; DB 4; Length 97;  
Best Local Similarity 100.0%; Pred. NO. 3e-06;  
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 TTNTPAMYVAIQAVL 15  
Db 5 TTNTPAMYVAIQAVL 19

RESULT 4  
AA26087  
ID AA26087 standard; protein; 97 AA.  
XX  
AC AA26087;  
XX  
DT 17-OCT-2001 (first entry)

XX Peptide #124 encoded by probe for measuring placental gene expression.  
XX

XX Probe; microarray; human; placenta; antenatal diagnosis;  
KW genetic disorder.

XX Homo sapiens.

XX WO200157272-A2.

XX 09-AUG-2001.

XX 30-JAN-2001; 2001WO-US000663.

XX 04-FEB-2000; 2000US-0180312P.

XX 26-MAY-2000; 2000US-0207456P.

XX 30-JUN-2000; 2000US-00608408.

XX 03-AUG-2000; 2000US-00632366.

XX 21-SEP-2000; 2000US-0234687P.

XX 27-SEP-2000; 2000US-0236359P.

XX 04-OCT-2000; 2000GB-00024263.

XX (MOLE-) MOLECULAR DYNAMICS INC.

XX Penn SG, Hanzel DK, Chen W, Rank DR;

XX WPI; 2001-488897/53.

XX Human genome-derived single exon nucleic acid probes useful for analyzing

XX gene expression in human placenta.

XX Claim 27; SEQ ID NO 26356; 654pp; English.

XX The present invention relates to single exon nucleic acid probes (SENPs;  
XX see AA131315-A157546). The present sequence is a peptide encoded by one  
XX such probe. The probes are useful for producing a microarray for  
XX predicting, measuring and displaying gene expression in samples derived  
XX from human placenta. The probes are useful for antenatal diagnosis of  
XX human genetic disorders

XX Sequence 97 AA;

Query Match 100.0%; Score 74; DB 4; Length 97;

Best Local Similarity 100.0%; Pred. No. 3e-06;

Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TTNTPAMVVAIQAVL 15

DB 5 TTNTPAMVVAIQAVL 19

RESULT 5

ABB27467

ID ABB27467 standard; peptide; 97 AA.

XX ABB27467;

XX 01-FEB-2002 (first entry)

XX Human peptide #118 encoded by breast cell single exon nucleic acid probe.

XX Human; microarray; single exon probe; gene expression; breast; disease;

XX cancer.

XX Homo sapiens.

XX WO200157271-A2.

XX 09-AUG-2001.

XX 30-JAN-2001; 2001WO-US000662.

XX 04-FEB-2000; 2000US-0180312P.

XX 26-MAY-2000; 2000US-0207456P.

XX 30-JUN-2000; 2000US-00608408.

PR 03-AUG-2000; 2000US-00632366.

PR 21-SEP-2000; 2000US-0234687P.

PR 27-SEP-2000; 2000US-0236359P.

PR 04-OCT-2000; 2000GB-00024263.

XX (MOLE-) MOLECULAR DYNAMICS INC.

XX Penn SG, Hanzel DK, Chen W, Rank DR;

XX WPI; 2001-496933/54.

XX New spatially-addressable set of single exon nucleic acid probes, useful  
XX for measuring gene expression in sample derived from human breast,  
XX comprises number of single exon nucleic acid probes.

XX Claim 27; SEQ ID NO 10435; 327pp + Sequence Listing; English.

XX The invention relates to a spatially-addressable set of single exon  
XX nucleic acid probes for measuring gene expression in a sample derived  
XX from human breast and BT 474 cells. The method involves contacting the  
XX probes with a collection of detectably labelled nucleic acids derived  
XX from mRNA of human breast, and then measuring the label bound to each  
XX probe of the microarray. The probes are useful for verifying the  
XX expression of regions of genomic DNA predicted to encode proteins. They  
XX are useful for gene discovery, and for determining predisposition and/or  
XX assessing the toxicity of chemical agents on cells. The microarray of  
XX this invention presents a far greater diversity of probes for measuring  
XX gene expression, with far less bias than expressed sequence tag  
XX microarrays. The method is suitable for rapid production of functional  
XX information from genomic sequence. The present sequence is a peptide  
XX encoded by a single exon nucleic acid probe of the invention. Note: The  
XX sequence data for this patent did not form part of the printed  
XX specification, but was obtained in electronic format directly from WIPO  
XX at ftp.wipo.int/pub/published\_pct\_sequences

XX Sequence 97 AA;

Query Match 100.0%; Score 74; DB 4; Length 97;

Best Local Similarity 100.0%; Pred. No. 3e-06;

Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TTNTPAMVVAIQAVL 15

DB 5 TTNTPAMVVAIQAVL 19

RESULT 6

ABB18116

ID ABB18116 standard; protein; 97 AA.

XX ABB18116;

XX 23-JAN-2002 (first entry)

XX Protein #115 encoded by probe for measuring heart cell gene expression.

XX Human; gene expression; heart; microarray; vascular system;

XX cardiovascular disease; hypertension; cardiac arrhythmia;

XX congenital heart disease.

XX Homo sapiens.

XX WO200157274-A2.

XX 09-AUG-2001.

XX 30-JAN-2001; 2001WO-US000666.

XX 04-FEB-2000; 2000US-0180312P.

XX 26-MAY-2000; 2000US-0207456P.

XX 30-JUN-2000; 2000US-00608408.

XX 03-AUG-2000; 2000US-00632366.



```
Query Match      100.0%; Score 74; DB 4; Length 97;
Best Local Similarity 100.0%; Pred. No. 3e-06;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 TFNTPAMYVAIQAVL 15
Db 5 TFNTPAMYVAIQAVL 19
|||||

RESULT 9
ABG47471
ID ABG47471 standard; peptide; 97 AA.
XX AC ABG47471;
XX DT 25-FEB-2003 (first entry)
XX DE Human liver peptide, SEQ ID No 26119.
XX KW Human; liver; cirrhosis; hyperlipoproteinaemia; hyperlipidaemia;
XX KW hypercholesterolaemia; coronary heart disease.
XX OS Homo sapiens.
XX PN WO200157273-A2.
XX PD 09-AUG-2001.
XX PF 30-JAN-2001; 2001WO-US000664.
XX PR 04-FEB-2000; 2000US-0180312P.
XX PR 26-MAY-2000; 2000US-0207456P.
XX PR 30-JUN-2000; 2000US-00608408.
XX PR 03-AUG-2000; 2000US-00632366.
XX PR 21-SEP-2000; 2000US-0234687P.
XX PR 27-SEP-2000; 2000US-0236359P.
XX PR 04-OCT-2000; 2000GB-00024263.
XX PA (MOLE-) MOLECULAR DYNAMICS INC.
XX PI Penn SG, Hanzel DK, Chen W, Rank DR;
XX WPI; 2001-488898/53.
XX DR Human genome-derived single exon nucleic acid probes useful for analyzing
XX PT gene expression in human adult liver.
XX PS Claim 27; SEQ ID NO 26119; 658pp; English.
XX CC The invention relates to a single exon nucleic acid probe (SENP) (I) for
XX CC measuring human gene expression in a sample derived from human adult
XX CC liver, comprising one of 13109 defined nucleotide sequences given in the
XX CC specification (or complements/ fragments). The probe hybridises at high
XX CC stringency to a nucleic acid molecule expressed in the human adult liver.
XX CC (I) may be used for predicting, measuring and displaying gene expression
XX CC in samples derived from human adult liver. The genes identified may be
XX CC involved in genetic liver diseases such as cirrhosis,
XX CC hyperlipoproteinaemia, hyperlipidaemia and hypercholesterolaemia which is
XX CC associated with coronary heart disease. ABG47348-ABG59930 represent human
XX CC liver single exon encoded peptides of the invention. Note: The sequence
XX CC information for this patent does not appear in the printed specification
XX CC but was obtained in electronic format directly from WIPO at
XX CC ftp.wipo.int/pub/published_pct_sequences
XX SQ Sequence 97 AA;

Query Match      100.0%; Score 74; DB 4; Length 97;
Best Local Similarity 100.0%; Pred. No. 3e-06;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 TFNTPAMYVAIQAVL 15
Db 5 TFNTPAMYVAIQAVL 19
|||||

RESULT 11
ABG35459
ID ABG35459 standard; peptide; 97 AA.
XX AC ABG35459;

Db 5 TFNTPAMYVAIQAVL 19
|||||

RESULT 10
AAM01437
ID AAM01437 standard; protein; 97 AA.
XX AC AAM01437;
XX DT 09-OCT-2001 (first entry)
XX DE Peptide #119 encoded by probe for measuring human breast gene expression.
XX DE Probe; human; breast disease; breast cancer; development disorder;
XX KW inflammatory disease; proliferative breast disease; non-carcinoma tumour.
XX OS Homo sapiens.
XX PN WO200157270-A2.
XX PD 09-AUG-2001.
XX PF 29-JAN-2001; 2001WO-US000661.
XX PR 04-FEB-2000; 2000US-0180312P.
XX PR 26-MAY-2000; 2000US-0207456P.
XX PR 30-JUN-2000; 2000US-00608408.
XX PR 03-AUG-2000; 2000US-00632366.
XX PR 21-SEP-2000; 2000US-0234687P.
XX PR 27-SEP-2000; 2000US-0236359P.
XX PR 04-OCT-2000; 2000GB-00024263.
XX PA (MOLE-) MOLECULAR DYNAMICS INC.
XX PI Penn SG, Hanzel DK, Chen W, Rank DR;
XX WPI; 2001-476286/51.
XX DR Novel single exon nucleic acid probe used to measuring gene expression in
XX PT a human breast.
XX PS Claim 27; SEQ ID NO 10177; 322pp; English.
XX CC The present invention relates to novel single exon nucleic acid probes
XX CC (see AAI00010-AAI10067). The present sequence is a peptide encoded by one
XX CC such probe. The probes are useful for measuring human gene expression in
XX CC a human breast sample, where the probe hybridises at high stringency to a
XX CC nucleic acid expressed in the human breast. The probes are useful for
XX CC predicting, diagnosing, grading, staging, monitoring and prognosing
XX CC diseases of the human breast, particularly those diseases with polygenic
XX CC aetiology. The diseases include: breast cancer, disorders of development,
XX CC inflammatory diseases of the breast, fibrocystic changes, proliferative
XX CC breast disease and non-carcinoma tumours. Note: The sequence data for
XX CC this patent did not form part of the printed specification, but was
XX CC obtained in electronic format directly from WIPO at
XX CC ftp.wipo.int/pub/published_pct_sequences
XX SQ Sequence 97 AA;

Query Match      100.0%; Score 74; DB 4; Length 97;
Best Local Similarity 100.0%; Pred. No. 3e-06;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 TFNTPAMYVAIQAVL 15
Db 5 TFNTPAMYVAIQAVL 19
|||||
```

19-AUG-2002 (first entry)  
Human peptide encoded by genome-derived single exon probe SEQ ID 25124.

Human; single exon probe; asthma; lung cancer; COPD; ILD;  
chronic obstructive pulmonary disease; interstitial lung disease;  
familial idiopathic pulmonary fibrosis; neurofibromatosis;  
tuberosus sclerosis; Gaucher's disease; Niemann-Pick disease;  
Hermansky-Pudlak syndrome; sarcoidosis; pulmonary haemosiderosis;  
pulmonary histiocytosis; lymphangioleiomyomatosis; Karagener syndrome;  
pulmonary alveolar proteinosis; fibrocystic pulmonary dysplasia;  
primary ciliary dyskinesia; pulmonary hypertension;  
hyaline membrane disease.

Homo sapiens.  
WO200186003-A2.  
15-NOV-2001.  
30-JAN-2001; 2001WO-US000665.  
04-FEB-2000; 2000US-0180312P.  
26-MAY-2000; 2000US-0207456P.  
30-JUN-2000; 2000US-00608408.  
03-AUG-2000; 2000US-00632366.  
21-SEP-2000; 2000US-0234687P.  
27-SEP-2000; 2000US-0236359P.  
04-OCT-2000; 2000GB-00024263.  
(MOLE-) MOLECULAR DYNAMICS INC.

Penn SG, Hanzel DK, Chen W, Rank DR;  
WPI; 2002-114183/15.

Spatially-addressable set of single exon nucleic acid probes, used to  
measure gene expression in human lung samples.

Claim 27; SEQ ID NO 25124; 634pp; English.

The invention relates to a spatially-addressable set of single exon  
nucleic acid probes for measuring gene expression in a sample derived  
from human lung comprising single exon nucleic acid probes having one of  
12614 nucleic acid sequences mentioned in the specification, or their  
complements or the 12387 open reading frames derived from the 12614  
probes. Also included are a microarray comprising the novel set of probes  
; the novel set of probes which hybridise at high stringency to a nucleic  
acid expressed in the human lung; measuring gene expression in a sample  
derived from human lung, comprising (a) contacting the array with a  
collection of detectably labeled nucleic acids derived from human lung  
mRNA, and (b) measuring the label detectably bound to each probe of the  
array; identifying exons in a eukaryotic genome, comprising (a)  
algorithmically predicting at least one exon from genomic sequences of  
the eukaryote; and (b) detecting specific hybridisation of detectably  
labeled nucleic acids from eukaryote lung mRNA, to a single exon probe,  
having a fragment identical to the predicted exon, the probe is included  
in the above mentioned microarray; assigning exons to a single gene,  
comprising (a) identifying exons from genomic sequence by the method  
above and (b) measuring the expression of each of the exons in several  
tissues and/or cell types using hybridisation to a single exon  
microarrays having a probe with the exon, where a common pattern of  
expression of the exons in the tissues and/or cell types indicates that  
the exons should be assigned to a single gene; a peptide comprising one  
of 12011 sequences, mentioned in the specification, or encoded by the  
probes/open reading frames (ORF). The probes are used for gene expression  
analysis, and for identifying exons in a gene, particularly using human  
lung derived mRNA and for the study of lung diseases such as asthma, lung  
cancer, chronic obstructive pulmonary disease (COPD), interstitial lung  
disease (ILD), familial idiopathic pulmonary fibrosis, neurofibromatosis,  
tuberosus sclerosis, Gaucher's disease, Niemann-Pick disease, Hermansky-  
Pudlak syndrome, sarcoidosis, pulmonary haemosiderosis, pulmonary

histiocytosis, lymphangioleiomyomatosis, pulmonary alveolar proteinosis,  
Karagener syndrome, fibrocystic pulmonary dysplasia, primary ciliary  
dyskinesia, pulmonary hypertension and hyaline membrane disease. The  
present sequence is a peptide/protein encoded by a single exon probe of  
the invention. Note: The sequence data for this patent did not form part  
of the printed specification, but was obtained in electronic format  
directly from WIPO at ftp.wipo.int/pub/published\_pct\_sequences

Sequence 97 AA;  
Query Match 100.0%; Score 74; DB 5; Length 97;  
Best Local Similarity 100.0%; Pred. No. 3e-06;  
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 T F T P A M Y V A I Q A V L 15  
D b | | | | | | | | | | | | | | |  
5 T F T P A M Y V A I Q A V L 19

RESULT 12  
AAU33076  
ID AAU33076 standard; protein; 137 AA.  
XX AC AAU33076;  
XX DT 18-DEC-2001 (first entry)  
XX DE Novel human secreted protein #3567.  
XX KW Human; vaccination; gene therapy; nutritional supplement;  
KW stem cell proliferation; haematopoiesis; nerve tissue regeneration;  
KW immune suppression; immune stimulation; anti-inflammatory; leukaemia.  
XX OS Homo sapiens.  
XX PN WO200179449-A2.  
XX PD 25-OCT-2001.  
XX PF 16-APR-2001; 2001WO-US008656.  
XX PR 18-APR-2000; 2000US-00552929.  
XX PR 26-JAN-2001; 2001US-00770160.  
XX PA (HYSE-) HYSEQ INC.  
XX FI Tang YT, Liu C, Drmanac RT;  
XX DR WPI; 2001-611725/70.  
XX PT Nucleic acids encoding a range of human polypeptides, useful in genetic  
PT vaccination, testing and therapy.  
XX PS Claim 20; Page 704-705; 765pp; English.  
XX CC The invention relates to novel human secreted polypeptides. The  
CC polypeptides and antibodies to the polypeptides are useful for  
CC determining the presence of or predisposition to a disease associated  
CC with altered levels of polypeptide. The polypeptides are also useful for  
CC identifying agents (agonists and antagonists) that bind to them. Cells  
CC for use in treatment of a pathology related to aberrant expression or  
CC physiological interactions of the polypeptide. Vectors comprising the  
CC nucleic acids encoding the polypeptides and cells genetically engineered  
CC to express them are also useful for producing the proteins. The proteins  
CC are useful in genetic vaccination, testing and therapy, and can be used  
CC as nutritional supplements. They may be used to increase stem cell  
CC proliferation; to regulate haematopoiesis; and in bone, cartilage, tendon  
CC and/or nerve tissue growth or regeneration; immune suppression and/or  
CC stimulation; as anti-inflammatory agents; and in treatment of leukaemias.  
CC AAU29510-AAU33304 represent the amino acid sequences of novel human  
CC secreted proteins of the invention  
XX





Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TTNTPAMYVAIQAVL 15  
 Db 137 TTNTPAMYVAIQAVL 151

RESULT 15  
 ABP42912  
 ID ABP42912 standard; protein; 168 AA.  
 XX  
 AC ABP42912;  
 XX  
 DT 22-AUG-2002 (first entry)  
 XX  
 DE Human ovarian antigen HPDQ265, SEQ ID NO:4044.  
 XX  
 KW Human; ovarian antigen; ovary; ovarian; breast; cancer; tumour;  
 KW ovarian cancer; breast cancer; tumour; reproductive system disorder;  
 KW infertility; pregnancy disorder; anovulation; polycystic ovary syndrome;  
 KW PCOS; ovarian cyst; dysmenorrhoea; endocrine disorder; infection;  
 KW inflammatory condition; immune disorder; blood disorder;  
 KW cardiovascular disorder; respiratory disorder; neurological disorder;  
 KW gastrointestinal disorder; urinary system disorder; drug screening;  
 KW gene therapy; chromosome mapping; forensic analysis;  
 KW antibody preparation; cytostatic; immunomodulatory; neuroprotective;  
 KW antiinflammatory; gynaecological; reproductive.  
 XX  
 OS Homo sapiens.  
 XX  
 PN WO200200677-A1.  
 XX  
 PD 03-JAN-2002.  
 XX  
 PF 07-JUN-2001; 2001WO-0018569.  
 XX  
 PR 07-JUN-2000; 2000US-0209467P.  
 XX  
 PA (HUMA-) HUMAN GENOME SCI INC.  
 XX  
 PI Birse CE, Rosen CA;  
 XX  
 DR WPI; 2002-147878/19.  
 DR N-PSDB; ABQ55989.  
 XX  
 PT Isolated nucleic acid molecules encoding novel ovarian polypeptides,  
 PT useful in the prevention, treatment and diagnosis of cancer (e.g. ovarian  
 PT cancer), immune disorders, cardiovascular disorders and neurological  
 PT diseases.  
 XX  
 PS Claim 11; SEQ ID NO 4044; 2922pp; English.  
 XX  
 CC The invention relates to 2175 novel human ovarian antigens (ABP41054-  
 CC ABP43228) and to cDNAs encoding them (ABQ54131-ABQ56305), and also  
 CC encompasses polypeptides 90% identical and polynucleotides 95% identical  
 CC to the sequences of the invention. The invention additionally relates to  
 CC recombinant vectors and host cells comprising human ovarian antigen  
 CC polynucleotides, antibodies against human ovarian antigens, and the use  
 CC of ovarian antigen polynucleotides and polypeptides in diagnosing,  
 CC treating, prognosing or preventing various ovary and/or breast-related  
 CC disorders. Such conditions include ovarian cancer and breast cancer, and  
 CC metastatic tumours of ovarian or breast origin, reproductive system  
 CC disorders (e.g., infertility, disorders of pregnancy, anovulation,  
 CC polycystic ovary syndrome, ovarian cysts, and dysmenorrhoea), endocrine  
 CC disorders, infections (e.g., chlamydia, HIV, toxoplasmosis, and toxic  
 CC shock syndrome), inflammatory conditions (e.g., mastitis, oophoritis and  
 CC vaginitis), immune disorders (e.g., congenital and acquired  
 CC immunodeficiencies, autoimmune oophoritis, systemic lupus erythematosus),  
 CC blood-related disorders (e.g., anaemia). Cardiovascular disorders  
 CC respiratory disorders, neurological disorders, gastrointestinal disorders  
 CC and urinary system disorders. Ovarian antigen polypeptides and  
 CC polynucleotides may also be used in screening for compounds which  
 CC modulate ovarian antigen expression or activity. The polynucleotides may

further be used for gene therapy, chromosome mapping, in the  
 identification of individuals and in forensic analysis, and the  
 polypeptides may be used as food additives or to prepare antibodies  
 useful in disease diagnosis, drug targeting and phenotyping. The present  
 sequence represents a human ovarian antigen of the invention. Note: The  
 sequence data for this patent did not form part of the printed  
 specification, but was obtained in electronic format directly from WIPO  
 at ftp.wipo.int/pub/published\_pct\_sequences

Sequence 168 AA;  
 Query Match 100.0%; Score 74; DB 5; Length 168;  
 Best Local Similarity 100.0%; Pred. No. 5.7e-06;  
 Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TTNTPAMYVAIQAVL 15  
 Db 12 TTNTPAMYVAIQAVL 26

Search completed: April 8, 2005, 10:50:45  
 Job time : 64.4 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: April 8, 2005, 09:58:57 ; Search time 17.5333 Seconds  
(without alignments)  
63.863 Million cell updates/sec

Title: US-09-423-351C-6  
Perfect score: 74  
Sequence: 1 TFPNTPMYVAIQAVL 15

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 513545 seqs, 74649064 residues

Total number of hits satisfying chosen parameters: 513545

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : Issued Patents AA:\*  
1: /cgn2\_6/ptodata/1/iaa/5A\_COMB.pep.\*  
2: /cgn2\_6/ptodata/1/iaa/5B\_COMB.pep.\*  
3: /cgn2\_6/ptodata/1/iaa/6A\_COMB.pep.\*  
4: /cgn2\_6/ptodata/1/iaa/6B\_COMB.pep.\*  
5: /cgn2\_6/ptodata/1/iaa/PCTUS\_COMB.pep.\*  
6: /cgn2\_6/ptodata/1/iaa/backfiles1.pep.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	74	100.0	374	3	US-08-609-236-6
2	74	100.0	375	4	US-09-976-594-731
3	74	100.0	399	4	US-09-949-016-9424
4	74	100.0	404	4	US-09-949-016-11313
5	69	93.2	375	3	US-09-106-217-16
6	69	93.2	376	1	US-08-588-113-2
7	69	93.2	377	3	US-09-949-016-6100
8	69	93.2	376	4	US-09-106-217-2
9	69	93.2	377	4	US-09-919-172-33
10	69	93.2	377	4	US-09-917-254-53
11	69	93.2	386	4	US-09-949-016-7721
12	69	93.2	402	4	US-09-949-016-10757
13	68	91.9	146	3	US-09-306-446C-5
14	68	91.9	374	3	US-09-306-446C-2
15	64	86.5	375	2	US-08-494-151-14
16	61	82.4	371	1	US-08-261-206A-77
17	61	82.4	375	3	US-09-171-337A-7
18	61	82.4	375	3	US-09-171-337A-8
19	61	82.4	375	4	US-09-631-022-7
20	61	82.4	375	4	US-09-631-022-8
21	61	82.4	445	4	US-09-248-796A-14109
22	56	75.7	447	4	US-09-949-016-8556
23	55	74.3	362	4	US-09-949-016-7725
24	55	74.3	376	4	US-09-538-092-1109
25	55	74.3	376	4	US-09-538-092-1110
26	55	74.3	376	4	US-09-949-016-6656
27	55	74.3	376	4	US-09-949-016-8452

28	51	68.9	336	4	US-09-248-796A-14108	Sequence 14108, A
29	47	63.5	365	4	US-09-248-796A-20619	Sequence 20619, A
30	42	56.8	394	4	US-09-949-016-6655	Sequence 6655, Ap
31	42	56.8	406	4	US-09-949-016-7396	Sequence 7396, Ap
32	40	54.1	620	3	US-09-232-200-49	Sequence 49, Appl
33	40	54.1	620	3	US-09-232-197-49	Sequence 49, Appl
34	40	54.1	620	3	US-09-232-201-49	Sequence 49, Appl
35	40	54.1	620	4	US-09-232-195-49	Sequence 49, Appl
36	40	54.1	687	4	US-09-949-016-10593	Sequence 10593, A
37	38	51.4	495	4	US-09-270-767-44726	Sequence 44726, A
38	38	51.4	620	3	US-09-232-200-61	Sequence 61, Appl
39	38	51.4	620	3	US-09-232-200-93	Sequence 93, Appl
40	38	51.4	620	3	US-09-232-197-61	Sequence 61, Appl
41	38	51.4	620	3	US-09-232-197-93	Sequence 93, Appl
42	38	51.4	620	3	US-09-232-201-61	Sequence 61, Appl
43	38	51.4	620	3	US-09-232-201-93	Sequence 93, Appl
44	38	51.4	620	4	US-09-232-195-61	Sequence 61, Appl
45	38	51.4	620	4	US-09-232-195-93	Sequence 93, Appl

## ALIGNMENTS

RESULT 1  
US-08-609-236-6  
; Sequence 6, Application US/08609236  
; Patent No. 6087398  
; GENERAL INFORMATION:  
; APPLICANT: Steven R. Goodman  
; TITLE OF INVENTION: NO. 6087398e1 Sickle Cell Anemia Treatment  
; NUMBER OF SEQUENCES: 6  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: McGregor & Adler, LLP  
; STREET: 8011 Candle Lane  
; CITY: Houston  
; STATE: Texas  
; COUNTRY: USA  
; ZIP: 77071  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: Apple Macintosh  
; OPERATING SYSTEM: Macintosh  
; SOFTWARE: Microsoft Word for Macintosh  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/609,236  
; FILING DATE: March 1, 1996  
; CLASSIFICATION: 514  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 60/002,288  
; FILING DATE: August 14, 1995  
; CLASSIFICATION: 514  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Benjamin Aaron Adler, Ph.D.  
; REGISTRATION NUMBER: 35,423  
; REFERENCE/DOCKET NUMBER: D5807  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 713-777-2321  
; TELEFAX: 713-777-6908  
; INFORMATION FOR SEQ ID NO: 6:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 374  
; TYPE: Amino acid  
; STRANDEDNESS:  
; TOPOLOGY: Linear  
; MOLECULE TYPE:  
; DESCRIPTION: Protein  
; HYPOTHETICAL: No  
; ANTI-SENSE: No  
; FRAGMENT TYPE:  
; ORIGINAL SOURCE:  
; STRAIN:  
; INDIVIDUAL ISOLATE:  
; DEVELOPMENTAL STAGE:

TISSUE TYPE:  
CELL TYPE:  
CELL LINE:  
US-08-609-236-6

Query Match 100.0%; Score 74; DB 3; Length 374;  
Best Local Similarity 100.0%; Pred. No. 1.3e-05;  
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TFTPAMYVAIQAVL 15  
Db 125 TFTPAMYVAIQAVL 139

RESULT 2  
US-09-976-594-731  
; Sequence 731, Application US/09976594  
; Patent No. 6673549  
; GENERAL INFORMATION:  
; APPLICANT: Furness, Michael  
; APPLICANT: Buchbinder, Jenny  
; TITLE OF INVENTION: GENES EXPRESSED IN C3A LIVER CELL CULTURES TREATED WITH STEROIDS  
; FILE REFERENCE: PA-0041 US  
; CURRENT APPLICATION NUMBER: US/09/976,594  
; CURRENT FILING DATE: 2001-10-12  
; PRIOR APPLICATION NUMBER: 60/240,409  
; PRIOR FILING DATE: 2000-10-12  
; NUMBER OF SEQ ID NOS: 1143  
; SOFTWARE: PERL Program  
; SEQ ID NO 731  
; LENGTH: 375  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
; NAME/KEY: misc feature  
; OTHER INFORMATION: Incyte ID No. 6673549 1837317CD1  
US-09-976-594-731

Query Match 100.0%; Score 74; DB 4; Length 375;  
Best Local Similarity 100.0%; Pred. No. 1.3e-05;  
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TFTPAMYVAIQAVL 15  
Db 126 TFTPAMYVAIQAVL 140

RESULT 3  
US-09-949-016-9424  
; Sequence 9424, Application US/09949016  
; Patent No. 6812339  
; GENERAL INFORMATION:  
; APPLICANT: VENTER, J. Craig et al.  
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED  
; TITLE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF  
; FILE REFERENCE: CL001307  
; CURRENT APPLICATION NUMBER: US/09/949,016  
; CURRENT FILING DATE: 2000-04-14  
; PRIOR APPLICATION NUMBER: 60/241,755  
; PRIOR FILING DATE: 2000-10-20  
; PRIOR APPLICATION NUMBER: 60/237,768  
; PRIOR FILING DATE: 2000-10-03  
; PRIOR APPLICATION NUMBER: 60/231,498  
; PRIOR FILING DATE: 2000-09-08  
; NUMBER OF SEQ ID NOS: 207012  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 9424  
; LENGTH: 399  
; TYPE: PRT  
; ORGANISM: Human  
US-09-949-016-9424

Query Match 100.0%; Score 74; DB 4; Length 399;

Best Local Similarity 100.0%; Pred. No. 1.4e-05;  
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TFTPAMYVAIQAVL 15  
Db 150 TFTPAMYVAIQAVL 164

RESULT 4  
US-09-949-016-11313  
; Sequence 11313, Application US/09949016  
; Patent No. 6812339  
; GENERAL INFORMATION:  
; APPLICANT: VENTER, J. Craig et al.  
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED  
; TITLE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF  
; FILE REFERENCE: CL001307  
; CURRENT APPLICATION NUMBER: US/09/949,016  
; CURRENT FILING DATE: 2000-04-14  
; PRIOR APPLICATION NUMBER: 60/241,755  
; PRIOR FILING DATE: 2000-10-20  
; PRIOR APPLICATION NUMBER: 60/237,768  
; PRIOR FILING DATE: 2000-10-03  
; PRIOR APPLICATION NUMBER: 60/231,498  
; PRIOR FILING DATE: 2000-09-08  
; NUMBER OF SEQ ID NOS: 207012  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 11313  
; LENGTH: 404  
; TYPE: PRT  
; ORGANISM: Human  
US-09-949-016-11313

Query Match 100.0%; Score 74; DB 4; Length 404;  
Best Local Similarity 100.0%; Pred. No. 1.4e-05;  
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TFTPAMYVAIQAVL 15  
Db 155 TFTPAMYVAIQAVL 169

RESULT 5  
US-09-106-217-16  
; Sequence 16, Application US/09106217  
; Patent No. 6063576  
; GENERAL INFORMATION:  
; APPLICANT: Keating, Mark T.  
; APPLICANT: Olson, Timothy M.  
; TITLE OF INVENTION: Actin Mutations in Dilated  
; TITLE OF INVENTION: Cardiomyopathy, a Heritable Form of Heart Failure  
; NUMBER OF SEQUENCES: 18  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Rothwell, Figg, Ernst & Kurz, P.C.  
; STREET: 555 Thirteenth Street, N.W., Suite 701 East  
; STREET: Tower  
; CITY: Washington  
; STATE: DC  
; COUNTRY: U.S.A.  
; ZIP: 20004  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patent In Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/09/106,217  
; FILING DATE:  
; CLASSIFICATION:  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Saxe, Stephen A.  
; REGISTRATION NUMBER: 38,609  
; REFERENCE/DOCKET NUMBER: 2323-125

US-09-423-351c-6.ra1

TISSUE TYPE:  
CELL TYPE:  
CELL LINE:  
US-08-609-236-6

Query Match 100.0%; Score 74; DB 3; Length 374;  
Best Local Similarity 100.0%; Pred. No. 1.3e-05;  
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TFTPAMYVAIQAVL 15  
Db 125 TFTPAMYVAIQAVL 139

RESULT 2  
US-09-976-594-731  
; Sequence 731, Application US/09976594  
; Patent No. 6673549  
; GENERAL INFORMATION:  
; APPLICANT: Furness, Michael  
; APPLICANT: Buchbinder, Jenny  
; TITLE OF INVENTION: GENES EXPRESSED IN C3A LIVER CELL CULTURES TREATED WITH STEROIDS  
; FILE REFERENCE: PA-0041 US  
; CURRENT APPLICATION NUMBER: US/09/976,594  
; CURRENT FILING DATE: 2001-10-12  
; PRIOR APPLICATION NUMBER: 60/240,409  
; PRIOR FILING DATE: 2000-10-12  
; NUMBER OF SEQ ID NOS: 1143  
; SOFTWARE: PERL Program  
; SEQ ID NO 731  
; LENGTH: 375  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
; NAME/KEY: misc feature  
; OTHER INFORMATION: Incyte ID No. 6673549 1837317CD1  
US-09-976-594-731

Query Match 100.0%; Score 74; DB 4; Length 375;  
Best Local Similarity 100.0%; Pred. No. 1.3e-05;  
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TFTPAMYVAIQAVL 15  
Db 126 TFTPAMYVAIQAVL 140

RESULT 3  
US-09-949-016-9424  
; Sequence 9424, Application US/09949016  
; Patent No. 6812339  
; GENERAL INFORMATION:  
; APPLICANT: VENTER, J. Craig et al.  
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED  
; TITLE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF  
; FILE REFERENCE: CL001307  
; CURRENT APPLICATION NUMBER: US/09/949,016  
; CURRENT FILING DATE: 2000-04-14  
; PRIOR APPLICATION NUMBER: 60/241,755  
; PRIOR FILING DATE: 2000-10-20  
; PRIOR APPLICATION NUMBER: 60/237,768  
; PRIOR FILING DATE: 2000-10-03  
; PRIOR APPLICATION NUMBER: 60/231,498  
; PRIOR FILING DATE: 2000-09-08  
; NUMBER OF SEQ ID NOS: 207012  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 9424  
; LENGTH: 399  
; TYPE: PRT  
; ORGANISM: Human  
US-09-949-016-9424

Query Match 100.0%; Score 74; DB 4; Length 399;

Best Local Similarity 100.0%; Pred. No. 1.4e-05;  
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TFTPAMYVAIQAVL 15  
Db 150 TFTPAMYVAIQAVL 164

RESULT 4  
US-09-949-016-11313  
; Sequence 11313, Application US/09949016  
; Patent No. 6812339  
; GENERAL INFORMATION:  
; APPLICANT: VENTER, J. Craig et al.  
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED  
; TITLE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF  
; FILE REFERENCE: CL001307  
; CURRENT APPLICATION NUMBER: US/09/949,016  
; CURRENT FILING DATE: 2000-04-14  
; PRIOR APPLICATION NUMBER: 60/241,755  
; PRIOR FILING DATE: 2000-10-20  
; PRIOR APPLICATION NUMBER: 60/237,768  
; PRIOR FILING DATE: 2000-10-03  
; PRIOR APPLICATION NUMBER: 60/231,498  
; PRIOR FILING DATE: 2000-09-08  
; NUMBER OF SEQ ID NOS: 207012  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 11313  
; LENGTH: 404  
; TYPE: PRT  
; ORGANISM: Human  
US-09-949-016-11313

Query Match 100.0%; Score 74; DB 4; Length 404;  
Best Local Similarity 100.0%; Pred. No. 1.4e-05;  
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TFTPAMYVAIQAVL 15  
Db 155 TFTPAMYVAIQAVL 169

RESULT 5  
US-09-106-217-16  
; Sequence 16, Application US/09106217  
; Patent No. 6063576  
; GENERAL INFORMATION:  
; APPLICANT: Keating, Mark T.  
; APPLICANT: Olson, Timothy M.  
; TITLE OF INVENTION: Actin Mutations in Dilated  
; TITLE OF INVENTION: Cardiomyopathy, a Heritable Form of Heart Failure  
; NUMBER OF SEQUENCES: 18  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Rothwell, Figg, Ernst & Kurz, P.C.  
; STREET: 555 Thirteenth Street, N.W., Suite 701 East  
; STREET: Tower  
; CITY: Washington  
; STATE: DC  
; COUNTRY: U.S.A.  
; ZIP: 20004  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patent In Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/09/106,217  
; FILING DATE:  
; CLASSIFICATION:  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Saxe, Stephen A.  
; REGISTRATION NUMBER: 38,609  
; REFERENCE/DOCKET NUMBER: 2323-125

TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 202-783-6040  
; TELEFAX: 202-783-6031  
; INFORMATION FOR SEQ ID NO: 16:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 375 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
US-09-106-217-16

Query Match 93.2%; Score 69; DB 3; Length 375;  
Best Local Similarity 93.3%; Pred. No. 0.00011;  
Matches 14; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 TTNTPAMYVAIQAVL 15  
Db 126 TTNVPMYVAIQAVL 140  
||| |||||

## RESULT 6

US-08-588-113-2  
; Sequence 2, Application US/08588113

Patent No. 5710003  
; GENERAL INFORMATION:  
; APPLICANT: McHugh, Kirk M.  
; TITLE OF INVENTION: DIAGNOSTIC TEST FOR DETERMINING  
; TITLE OF INVENTION: MALIGNANCY OF SMOOTH MUSCLE TUMORS  
; NUMBER OF SEQUENCES: 4  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Woodcock Washburn Kurtz Mackiewicz &  
; ADDRESSEE: No. 5710003ris  
; STREET: One Liberty Place - 46th Floor  
; CITY: Philadelphia  
; STATE: PA  
; COUNTRY: USA  
; ZIP: 19103

COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patent In Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/588,113  
; FILING DATE:  
; CLASSIFICATION: 436

ATTORNEY/AGENT INFORMATION:  
; NAME: Ralph, Rebecca L.  
; REGISTRATION NUMBER: 35,152  
; REFERENCE/DOCKET NUMBER: TJU-1652  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 215-568-3100  
; TELEFAX: 215-568-3439  
; INFORMATION FOR SEQ ID NO: 2:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 376 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
US-08-588-113-2

Query Match 93.2%; Score 69; DB 1; Length 376;  
Best Local Similarity 93.3%; Pred. No. 0.00011;  
Matches 14; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 TTNTPAMYVAIQAVL 15  
Db 127 TTNVPMYVAIQAVL 141  
||| |||||

## RESULT 7

US-09-949-016-6100  
; Sequence 6100, Application US/09949016

Patent No. 6812339  
; GENERAL INFORMATION:  
; APPLICANT: VENTER, J. Craig et al.  
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED  
; TITLE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF  
; FILE REFERENCE: CL001307  
; CURRENT APPLICATION NUMBER: US/09/949,016  
; CURRENT FILING DATE: 2000-04-14  
; PRIOR APPLICATION NUMBER: 60/241,755  
; PRIOR FILING DATE: 2000-10-20  
; PRIOR APPLICATION NUMBER: 60/237,768  
; PRIOR FILING DATE: 2000-10-03  
; PRIOR APPLICATION NUMBER: 60/231,498  
; PRIOR FILING DATE: 2000-09-08  
; NUMBER OF SEQ ID NOS: 207012  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 6100  
; LENGTH: 376  
; TYPE: PRT  
; ORGANISM: Human  
US-09-949-016-6100

Query Match 93.2%; Score 69; DB 4; Length 376;  
Best Local Similarity 93.3%; Pred. No. 0.00011;  
Matches 14; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 TTNTPAMYVAIQAVL 15  
Db 127 TTNVPMYVAIQAVL 141  
||| |||||

## RESULT 8

US-09-106-217-2  
; Sequence 2, Application US/09106217

Patent No. 6063576  
; GENERAL INFORMATION:  
; APPLICANT: Keating, Mark T.  
; APPLICANT: Olson, Timothy M.  
; TITLE OF INVENTION: Actin Mutations in Dilated  
; TITLE OF INVENTION: Cardiomyopathy, a Heritable Form of Heart Failure  
; NUMBER OF SEQUENCES: 18  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Rothwell, Figg, Ernst & Kurz, P.C.  
; STREET: 555 Thirteenth Street, N.W., Suite 701 East  
; STREET: Tower  
; CITY: Washington  
; STATE: DC  
; COUNTRY: U.S.A.  
; ZIP: 20004

COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patent In Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/09/106,217  
; FILING DATE:  
; CLASSIFICATION:

ATTORNEY/AGENT INFORMATION:  
; NAME: Saxe, Stephen A.  
; REGISTRATION NUMBER: 38,609  
; REFERENCE/DOCKET NUMBER: 2323-125  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 202-783-6040  
; TELEFAX: 202-783-6031  
; INFORMATION FOR SEQ ID NO: 2:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 377 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
US-09-106-217-2

```
Query Match      93.2%; Score 69; DB 3; Length 377;
Best Local Similarity 93.3%; Pred. No. 0.00011;
Matches 14; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY      1 TTNTPAMYVAIQAVL 15
Db      128 TTNVPAMYVAIQAVL 142

RESULT 9
US-09-919-172-33
; Sequence 33, Application US/09919172
; Patent No. 6673545
; GENERAL INFORMATION:
; APPLICANT: Farris, Mary
; TITLE OF INVENTION: PROSTATE CANCER MARKERS
; FILE REFERENCE: PA-0036 US
; CURRENT APPLICATION NUMBER: US/09/919,172
; CURRENT FILING DATE: 2001-07-30
; PRIOR APPLICATION NUMBER: 60/222,469
; PRIOR FILING DATE: 2000-07-28
; NUMBER OF SEQ ID NOS: 102
; SOFTWARE: PERL Program
; SEQ ID NO 33
; LENGTH: 377
; TYPE: PRT
; ORGANISM: Homo sapiens
; NAME/KEY: misc feature
; OTHER INFORMATION: Incyte ID No. 6673545 1709118CD1
US-09-919-172-33

Query Match      93.2%; Score 69; DB 4; Length 377;
Best Local Similarity 93.3%; Pred. No. 0.00011;
Matches 14; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY      1 TTNTPAMYVAIQAVL 15
Db      128 TTNVPAMYVAIQAVL 142

RESULT 10
US-09-917-254-53
; Sequence 53, Application US/09917254
; Patent No. 6703204
; GENERAL INFORMATION:
; APPLICANT: Mutter, George
; TITLE OF INVENTION: Prognostic Classification of Breast Cancer
; FILE REFERENCE: B0801/7224(JRV)
; CURRENT APPLICATION NUMBER: US/09/917,254
; CURRENT FILING DATE: 2001-07-27
; PRIOR APPLICATION NUMBER: US 60/222,093
; PRIOR FILING DATE: 2000-07-28
; NUMBER OF SEQ ID NOS: 102
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 53
; LENGTH: 377
; TYPE: PRT
; ORGANISM: Homo Sapiens
US-09-917-254-53

Query Match      93.2%; Score 69; DB 4; Length 377;
Best Local Similarity 93.3%; Pred. No. 0.00011;
Matches 14; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY      1 TTNTPAMYVAIQAVL 15
Db      128 TTNVPAMYVAIQAVL 142

RESULT 11
US-09-919-172-33
; Sequence 33, Application US/09919172
; Patent No. 6673545
; GENERAL INFORMATION:
; APPLICANT: Farris, Mary
; TITLE OF INVENTION: PROSTATE CANCER MARKERS
; FILE REFERENCE: PA-0036 US
; CURRENT APPLICATION NUMBER: US/09/919,172
; CURRENT FILING DATE: 2001-07-30
; PRIOR APPLICATION NUMBER: 60/222,469
; PRIOR FILING DATE: 2000-07-28
; NUMBER OF SEQ ID NOS: 102
; SOFTWARE: PERL Program
; SEQ ID NO 33
; LENGTH: 377
; TYPE: PRT
; ORGANISM: Homo sapiens
; NAME/KEY: misc feature
; OTHER INFORMATION: Incyte ID No. 6673545 1709118CD1
US-09-919-172-33

Query Match      93.2%; Score 69; DB 3; Length 377;
Best Local Similarity 93.3%; Pred. No. 0.00011;
Matches 14; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY      1 TTNTPAMYVAIQAVL 15
Db      128 TTNVPAMYVAIQAVL 142

RESULT 12
US-09-949-016-10757
; Sequence 7721, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 7721
; LENGTH: 386
; TYPE: PRT
; ORGANISM: Human
US-09-949-016-10757

Query Match      93.2%; Score 69; DB 4; Length 386;
Best Local Similarity 93.3%; Pred. No. 0.00011;
Matches 14; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY      1 TTNTPAMYVAIQAVL 15
Db      137 TTNVPAMYVAIQAVL 151

RESULT 13
US-09-949-016-10757
; Sequence 7721, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 10757
; LENGTH: 402
; TYPE: PRT
; ORGANISM: Human
US-09-949-016-10757

Query Match      93.2%; Score 69; DB 4; Length 402;
Best Local Similarity 93.3%; Pred. No. 0.00012;
Matches 14; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY      1 TTNTPAMYVAIQAVL 15
Db      153 TTNVPAMYVAIQAVL 167

RESULT 13
US-09-306-446C-5
; Sequence 5, Application US/09306446C
; Patent No. 6372959
; GENERAL INFORMATION:
; APPLICANT: KIM, Dong Soo
```

```

; APPLICANT: KIM, Chul Geun
; APPLICANT: NAM, Yoon Kwon
; APPLICANT: NOH, Jae Koo
; APPLICANT: CHO, Kyou Nam
; TITLE OF INVENTION: EXPRESSION VECTOR OF MUD LOACH GROWTH HORMONE GENE
; FILE REFERENCE: P06344USO/BAS
; CURRENT APPLICATION NUMBER: US/09/306.446C
; CURRENT FILING DATE: 1999-05-06
; PRIOR APPLICATION NUMBER: KR 98/20255
; PRIOR FILING DATE: 1998-06-01
; NUMBER OF SEQ ID NOS: 33
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 5
; LENGTH: 146
; TYPE: PRT
; ORGANISM: Misgurnus mizolepus
US-09-306-446C-5

Query Match 91.9%; Score 68; DB 3; Length 146;
Best Local Similarity 93.3%; Pred. No. 5.6e-05;
Matches 14; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

```

```

Qy 1 TFNTPAMYVAIQAVL 15
| | | | | | | | | | | | | | | |
Db 5 TFNTPAMYVAIQAVL 19

```

```

RESULT 14
US-09-306-446C-2
; Sequence 2, Application US/09306446C
; Patent No. 6372959
; GENERAL INFORMATION:
; APPLICANT: KIM, Dong Soo
; APPLICANT: KIM, Chul Geun
; APPLICANT: NAM, Yoon Kwon
; APPLICANT: NOH, Jae Koo
; APPLICANT: CHO, Kyou Nam
; TITLE OF INVENTION: EXPRESSION VECTOR OF MUD LOACH GROWTH HORMONE GENE
; FILE REFERENCE: P06344USO/BAS
; CURRENT APPLICATION NUMBER: US/09/306.446C
; CURRENT FILING DATE: 1999-05-06
; PRIOR APPLICATION NUMBER: KR 98/20255
; PRIOR FILING DATE: 1998-06-01
; NUMBER OF SEQ ID NOS: 33
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 2
; LENGTH: 374
; TYPE: PRT
; ORGANISM: Misgurnus mizolepus
US-09-306-446C-2

```

```

Query Match 91.9%; Score 68; DB 3; Length 374;
Best Local Similarity 93.3%; Pred. No. 0.00017;
Matches 14; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

```

```

Qy 1 TFNTPAMYVAIQAVL 15
| | | | | | | | | | | | | | | |
Db 126 TFNTPAMYVAIQAVL 140

```

```

RESULT 15
US-08-494-151-14
; Sequence 14, Application US/08494151
; Patent No. 5840528
; GENERAL INFORMATION:
; APPLICANT: Van Ooyen, Albert Johannes Joseph
; TITLE OF INVENTION: Transformation of Phaffia rhodozyma
; NUMBER OF SEQUENCES: 14
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Morrison & Foerster
; STREET: 2000 Pennsylvania Ave. N.W., Suite 5500
; CITY: Washington, D.C.
; COUNTRY: USA

```

```

; ZIP: 20006-1812
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/494,151
; FILING DATE: 23-JUN-1995
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Murashige, Kate H.
; REGISTRATION NUMBER: 29,959
; REFERENCE/DOCKET NUMBER: 24615-20039.01
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202) 887-1500
; TELEFAX: (202) 887-0763
; TELEX: 90-4030
; INFORMATION FOR SEQ ID NO: 14:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 375 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-494-151-14

```

```

Query Match 86.5%; Score 64; DB 2; Length 375;
Best Local Similarity 86.7%; Pred. No. 0.00092;
Matches 13; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

```

```

Qy 1 TFNTPAMYVAIQAVL 15
| | | | | | | | | | | | | | | |
Db 126 TFNAPAFYVAIQAVL 140

```

```

Search completed: April 8, 2005, 12:07:36
Job time : 18.5333 secs

```

**THIS PAGE BLANK (USPTO)**



GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: April 8, 2005, 10:53:18 ; Search time 42.6667 Seconds  
(without alignments)  
116.718 Million cell updates/sec

Title: US-09-423-351C-6

Perfect score: 74  
Sequence: 1 TFNTPAMVIAQAVL 15

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 1418010 seqs, 331997259 residues

Total number of hits satisfying chosen parameters: 1418010

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : Published Applications AA:\*  
1: /cgn2\_6/ptodata/1/pubpaa/US07\_PUBCOMB.pep.\*  
2: /cgn2\_6/ptodata/1/pubpaa/PCT\_NEW\_PUB.pep.\*  
3: /cgn2\_6/ptodata/1/pubpaa/US05\_NEW\_PUB.pep.\*  
4: /cgn2\_6/ptodata/1/pubpaa/US06\_PUBCOMB.pep.\*  
5: /cgn2\_6/ptodata/1/pubpaa/US07\_NEW\_PUB.pep.\*  
6: /cgn2\_6/ptodata/1/pubpaa/PCTUS\_PUBCOMB.pep.\*  
7: /cgn2\_6/ptodata/1/pubpaa/US08\_NEW\_PUB.pep.\*  
8: /cgn2\_6/ptodata/1/pubpaa/US08\_PUBCOMB.pep.\*  
9: /cgn2\_6/ptodata/1/pubpaa/US09A\_PUBCOMB.pep.\*  
10: /cgn2\_6/ptodata/1/pubpaa/US09B\_PUBCOMB.pep.\*  
11: /cgn2\_6/ptodata/1/pubpaa/US09C\_PUBCOMB.pep.\*  
12: /cgn2\_6/ptodata/1/pubpaa/US09\_NEW\_PUB.pep.\*  
13: /cgn2\_6/ptodata/1/pubpaa/US10A\_PUBCOMB.pep.\*  
14: /cgn2\_6/ptodata/1/pubpaa/US10B\_PUBCOMB.pep.\*  
15: /cgn2\_6/ptodata/1/pubpaa/US10C\_PUBCOMB.pep.\*  
16: /cgn2\_6/ptodata/1/pubpaa/US10D\_PUBCOMB.pep.\*  
17: /cgn2\_6/ptodata/1/pubpaa/US10\_NEW\_PUB.pep.\*  
18: /cgn2\_6/ptodata/1/pubpaa/US11\_NEW\_PUB.pep.\*  
19: /cgn2\_6/ptodata/1/pubpaa/US60\_NEW\_PUB.pep.\*  
20: /cgn2\_6/ptodata/1/pubpaa/US60\_PUBCOMB.pep.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Match	Length	ID	Description
1	74	100.0	97	US-09-864-761-33414	Sequence 33414, A
2	74	100.0	168	US-10-264-049-4044	Sequence 4044, Ap
3	74	100.0	180	US-10-437-963-169247	Sequence 169247, A
4	74	100.0	197	US-10-424-599-280810	Sequence 280810, A
5	74	100.0	204	US-10-264-049-4148	Sequence 4148, Ap
6	74	100.0	208	US-10-424-599-282905	Sequence 282905, A
7	74	100.0	219	US-10-002-631C-56	Sequence 56, Appl
8	74	100.0	237	US-10-424-599-143226	Sequence 143226, A
9	74	100.0	261	US-10-424-599-177530	Sequence 177530, A
10	74	100.0	276	US-10-425-114-52089	Sequence 52089, A
11	74	100.0	299	US-10-424-599-278431	Sequence 278431, A
12	74	100.0	371	US-10-322-281-68	Sequence 68, Appl
13	74	100.0	375	US-10-205-194-93	Sequence 93, Appl

14	74	100.0	375	14	US-10-316-253-88	Sequence 88, Appl
15	74	100.0	375	15	US-10-205-331-94	Sequence 94, Appl
16	74	100.0	375	15	US-10-260-708-82	Sequence 82, Appl
17	74	100.0	377	15	US-10-424-599-223492	Sequence 223492, A
18	74	100.0	379	15	US-10-424-599-278432	Sequence 278432, A
19	74	100.0	380	16	US-10-437-963-187415	Sequence 187415, A
20	74	100.0	381	15	US-10-424-599-280814	Sequence 280814, A
21	74	100.0	398	16	US-10-322-281-63	Sequence 63, Appl
22	74	100.0	413	9	US-09-925-301-1436	Sequence 1436, Ap
23	70	94.6	183	15	US-10-264-049-3432	Sequence 3432, Ap
24	70	94.6	378	15	US-10-425-114-42317	Sequence 42317, A
25	69	93.2	108	9	US-09-864-761-37942	Sequence 37942, A
26	69	93.2	248	15	US-10-424-599-162684	Sequence 162684, A
27	69	93.2	250	15	US-10-425-114-44494	Sequence 44494, A
28	69	93.2	254	15	US-10-425-114-69353	Sequence 69353, A
29	69	93.2	288	15	US-10-425-114-57500	Sequence 57500, A
30	69	93.2	331	15	US-10-425-114-59172	Sequence 59172, A
31	69	93.2	347	15	US-10-424-599-283333	Sequence 283333, A
32	69	93.2	376	15	US-10-341-434-101	Sequence 101, App
33	69	93.2	377	9	US-09-919-172-33	Sequence 33, Appl
34	69	93.2	377	14	US-10-338-777-52	Sequence 52, Appl
35	69	93.2	377	15	US-10-236-031B-64	Sequence 64, Appl
36	69	93.2	377	15	US-10-424-599-162009	Sequence 162009, A
37	69	93.2	377	15	US-10-424-599-162685	Sequence 162685, A
38	69	93.2	377	15	US-10-424-599-162686	Sequence 162686, A
39	69	93.2	377	15	US-10-424-599-283336	Sequence 283336, A
40	69	93.2	377	16	US-10-333-680-4	Sequence 4, Appli
41	69	93.2	377	16	US-10-437-963-148877	Sequence 148877, A
42	69	93.2	377	16	US-10-437-963-198295	Sequence 198295, A
43	69	93.2	377	16	US-10-767-701-47239	Sequence 47239, A
44	69	93.2	378	15	US-10-425-114-52458	Sequence 52458, A
45	69	93.2	378	15	US-10-425-114-59265	Sequence 59265, A

## ALIGNMENTS

RESULT 1  
US-09-864-761-33414  
; Sequence 33414, Application US/09864761  
; Patent No. US20020048763A1  
; GENERAL INFORMATION:  
; APPLICANT: Penn, Sharron G.  
; APPLICANT: Rank, David R.  
; APPLICANT: Hanzel, David K.  
; APPLICANT: Chen, Wensheng  
; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR  
; FILE REFERENCE: Aecmica-X-1  
; CURRENT APPLICATION NUMBER: US/09/864,761  
; CURRENT FILING DATE: 2001-05-23  
; PRIOR APPLICATION NUMBER: US 60/180,312  
; PRIOR FILING DATE: 2000-02-04  
; PRIOR APPLICATION NUMBER: US 60/207,456  
; PRIOR FILING DATE: 2000-05-26  
; PRIOR APPLICATION NUMBER: US 09/632,366  
; PRIOR FILING DATE: 2000-08-03  
; PRIOR APPLICATION NUMBER: GB 24263.6  
; PRIOR FILING DATE: 2000-10-04  
; PRIOR APPLICATION NUMBER: US 60/236,359  
; PRIOR FILING DATE: 2000-09-27  
; PRIOR APPLICATION NUMBER: PCT/US01/00666  
; PRIOR FILING DATE: 2001-01-30  
; PRIOR APPLICATION NUMBER: PCT/US01/00667  
; PRIOR FILING DATE: 2001-01-30  
; PRIOR APPLICATION NUMBER: PCT/US01/00664  
; PRIOR FILING DATE: 2001-01-30  
; PRIOR APPLICATION NUMBER: PCT/US01/00669  
; PRIOR FILING DATE: 2001-01-30  
; PRIOR APPLICATION NUMBER: PCT/US01/00665  
; PRIOR FILING DATE: 2001-01-30  
; PRIOR APPLICATION NUMBER: PCT/US01/00668  
; PRIOR FILING DATE: 2001-01-30

; PRIOR APPLICATION NUMBER: PCT/US01/00663  
; PRIOR FILING DATE: 2001-01-30  
; PRIOR APPLICATION NUMBER: PCT/US01/00662  
; PRIOR FILING DATE: 2001-01-30  
; PRIOR APPLICATION NUMBER: PCT/US01/00661  
; PRIOR FILING DATE: 2001-01-30  
; PRIOR APPLICATION NUMBER: PCT/US01/00670  
; PRIOR FILING DATE: 2001-01-30  
; PRIOR APPLICATION NUMBER: US 60/234,687  
; PRIOR FILING DATE: 2000-09-21  
; PRIOR APPLICATION NUMBER: US 09/608,408  
; PRIOR FILING DATE: 2000-06-30  
; PRIOR APPLICATION NUMBER: US 09/774,203  
; PRIOR FILING DATE: 2001-01-29  
; NUMBER OF SEQ ID NOS: 49117  
; SOFTWARE: Annomax Sequence Listing Engine vers. 1.1  
; SEQ ID NO 33414  
; LENGTH: 97  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
; FEATURE:  
; OTHER INFORMATION: MAP TO AC006483.2  
; OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 18  
; OTHER INFORMATION: EXPRESSED IN HEART, SIGNAL = 48  
; OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 81  
; OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 28  
; OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 24  
; OTHER INFORMATION: EXPRESSED IN BT474, SIGNAL = 85  
; OTHER INFORMATION: EXPRESSED IN HBL100, SIGNAL = 78  
; OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 37  
; OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 45  
; OTHER INFORMATION: EXPRESSED IN HELA, SIGNAL = 40  
; OTHER INFORMATION: EST\_HUMAN HIT: BE271730.1, EVALUE 3.00e-50  
; OTHER INFORMATION: SWISSPROT HIT: O42161, EVALUE 4.00e-51  
; OTHER INFORMATION: EST\_HUMAN HIT: BE256272.1, EVALUE 4.00e-50

US-09-864-761-33414

Query Match 100.0%; Score 74; DB 9; Length 97;  
Best Local Similarity 100.0%; Pred. No. 9.4e-06; Mismatches 0; Indels 0; Gaps 0;  
Matches 15; Conservative 0;

Qy 1 TFNTPAMYVAIQAVL 15  
Db 5 TFNTPAMYVAIQAVL 19

RESULT 2  
US-10-264-049-4044  
; Sequence 4044, Application US/10264049  
; Publication No. US20040005579A1  
; GENERAL INFORMATION:  
; APPLICANT: Bires et al.  
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies  
; FILE REFERENCE: PA133P1  
; CURRENT APPLICATION NUMBER: US/10/264,049  
; CURRENT FILING DATE: 2002-10-04  
; PRIOR APPLICATION NUMBER: PCT/US01/18569  
; PRIOR FILING DATE: 2001-06-07  
; PRIOR APPLICATION NUMBER: US 60/209,467  
; PRIOR FILING DATE: 2000-06-07  
; NUMBER OF SEQ ID NOS: 4360  
; SOFTWARE: PatentIn Ver. 3.1  
; SEQ ID NO 4044  
; LENGTH: 168  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-10-264-049-4044

Query Match 100.0%; Score 74; DB 15; Length 168;  
Best Local Similarity 100.0%; Pred. No. 1.7e-05; Mismatches 0; Indels 0; Gaps 0;  
Matches 15; Conservative 0;

Qy 1 TFNTPAMYVAIQAVL 15

Db 12 TFNTPAMYVAIQAVL 26

RESULT 3  
US-10-437-963-169247  
; Sequence 169247, Application US/10437963  
; Publication No. US20040123343A1  
; GENERAL INFORMATION:  
; APPLICANT: La Rosa, Thomas J.  
; APPLICANT: Kovalic, David K.  
; APPLICANT: Zhou, Yihua  
; APPLICANT: Cao, Yongwei  
; APPLICANT: Wu, Wei  
; APPLICANT: Boukharov, Andrey A.  
; APPLICANT: Barbazuk, Brad  
; APPLICANT: Li, Ping  
; TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated With  
; FILE REFERENCE: 38-21(53221)B  
; CURRENT APPLICATION NUMBER: US/10/437,963  
; CURRENT FILING DATE: 2003-05-14  
; NUMBER OF SEQ ID NOS: 204966  
; SEQ ID NO 169247  
; LENGTH: 180  
; TYPE: PRT  
; ORGANISM: Oryza sativa  
; FEATURE:  
; OTHER INFORMATION: Clone ID: PAT\_MRT4530\_67685C.1.pep  
US-10-437-963-169247

Query Match 100.0%; Score 74; DB 16; Length 180;  
Best Local Similarity 100.0%; Pred. No. 1.9e-05; Mismatches 0; Indels 0; Gaps 0;  
Matches 15; Conservative 0;

Qy 1 TFNTPAMYVAIQAVL 15  
Db 61 TFNTPAMYVAIQAVL 75

RESULT 4  
US-10-424-599-280810  
; Sequence 280810, Application US/10424599  
; Publication No. US20040031072A1  
; GENERAL INFORMATION:  
; APPLICANT: La Rosa Thomas J  
; APPLICANT: Kovalic David K  
; APPLICANT: Zhou Yihua  
; APPLICANT: Cao Yongwei  
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With  
; FILE REFERENCE: 38-21(53223)B  
; CURRENT APPLICATION NUMBER: US/10/424,599  
; CURRENT FILING DATE: 2003-04-28  
; NUMBER OF SEQ ID NOS: 285684  
; SEQ ID NO 280810  
; LENGTH: 197  
; TYPE: PRT  
; ORGANISM: Glycine max  
; FEATURE:  
; NAME/KEY: unsure  
; LOCATION: (1)..(197)  
; OTHER INFORMATION: unsure at all Xaa locations  
; FEATURE:  
; OTHER INFORMATION: Clone ID: PAT\_MRT3847\_95594C.1.pep  
US-10-424-599-280810

Query Match 100.0%; Score 74; DB 15; Length 197;  
Best Local Similarity 100.0%; Pred. No. 2.1e-05; Mismatches 0; Indels 0; Gaps 0;  
Matches 15; Conservative 0;

Qy 1 TFNTPAMYVAIQAVL 15

```
Db      128  TFTPAMYVAIQAVL 142
; Publication No. US20030157486A1
; GENERAL INFORMATION:
; APPLICANT: Graff, Jonathon M.
; APPLICANT: Muenster, Matthew
; TITLE OF INVENTION: METHODS TO IDENTIFY SIGNAL SEQUENCES
; FILE REFERENCE: A34943 090495.0243
; CURRENT APPLICATION NUMBER: US/10/002.631C
; CURRENT FILING DATE: 2001-10-31
; PRIOR APPLICATION NUMBER: 60/300,309
; PRIOR FILING DATE: 2001-06-21
; NUMBER OF SEQ ID NOS: 324
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 56
; LENGTH: 219
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-002-631C-56

Query Match      100.0%; Score 74; DB 14; Length 219;
Best Local Similarity 100.0%; Pred. No. 2.3e-05;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1  TFTPAMYVAIQAVL 15
      |||||||
DB      153  TFTPAMYVAIQAVL 167

RESULT 8
US-10-424-599-143226
; Sequence 143226, Application US/10424599
; Publication No. US20040031072A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa Thomas J
; APPLICANT: Kovalic David K
; APPLICANT: Zhou Yihua
; APPLICANT: Cao Yongwei
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(53223)B
; CURRENT APPLICATION NUMBER: US/10/424,599
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 285684
; SEQ ID NO 143226
; LENGTH: 237
; TYPE: PRT
; ORGANISM: Glycine max
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (1)..(237)
; OTHER INFORMATION: unsure at all Xaa locations
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT3847_100347C.1.pep
US-10-424-599-143226

Query Match      100.0%; Score 74; DB 15; Length 237;
Best Local Similarity 100.0%; Pred. No. 2.5e-05;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1  TFTPAMYVAIQAVL 15
      |||||||
DB      10  TFTPAMYVAIQAVL 24

RESULT 9
US-10-424-599-177530
; Sequence 177530, Application US/10424599
; Publication No. US20040031072A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa Thomas J
; APPLICANT: Kovalic David K
; APPLICANT: Zhou Yihua
; APPLICANT: Cao Yongwei
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With
```

---

```
Db      128  TFTPAMYVAIQAVL 142
; Publication No. US20040005579A1
; GENERAL INFORMATION:
; APPLICANT: Birse et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
; FILE REFERENCE: PA133P1
; CURRENT APPLICATION NUMBER: US/10/264,049
; CURRENT FILING DATE: 2002-10-04
; PRIOR APPLICATION NUMBER: PCT/US01/18569
; PRIOR FILING DATE: 2001-06-07
; PRIOR APPLICATION NUMBER: US 60/209,467
; PRIOR FILING DATE: 2000-06-07
; NUMBER OF SEQ ID NOS: 4360
; SOFTWARE: PatentIn Ver. 3.1
; SEQ ID NO 4148
; LENGTH: 204
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-264-049-4148

Query Match      100.0%; Score 74; DB 15; Length 204;
Best Local Similarity 100.0%; Pred. No. 2.1e-05;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1  TFTPAMYVAIQAVL 15
      |||||||
DB      146  TFTPAMYVAIQAVL 160

RESULT 6
US-10-424-599-282905
; Sequence 282905, Application US/10424599
; Publication No. US20040031072A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa Thomas J
; APPLICANT: Kovalic David K
; APPLICANT: Zhou Yihua
; APPLICANT: Cao Yongwei
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(53223)B
; CURRENT APPLICATION NUMBER: US/10/424,599
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 285684
; SEQ ID NO 282905
; LENGTH: 208
; TYPE: PRT
; ORGANISM: Glycine max
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (1)..(208)
; OTHER INFORMATION: unsure at all Xaa locations
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT3847_97486C.1.pep
US-10-424-599-282905

Query Match      100.0%; Score 74; DB 15; Length 208;
Best Local Similarity 100.0%; Pred. No. 2.2e-05;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1  TFTPAMYVAIQAVL 15
      |||||||
DB      128  TFTPAMYVAIQAVL 142

RESULT 7
US-10-002-631C-56
; Sequence 56, Application US/10002631C
```

; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement  
; FILE REFERENCE: 38-21(53223)B  
; CURRENT APPLICATION NUMBER: US/10/424,599  
; CURRENT FILING DATE: 2003-04-28  
; NUMBER OF SEQ ID NOS: 285684  
; SEQ ID NO 177530  
; LENGTH: 261  
; TYPE: PRT  
; ORGANISM: Glycine max  
; FEATURE:  
; OTHER INFORMATION: Clone ID: PAT\_MRT3847\_131325C.1.pep  
US-10-424-599-177530

Query Match 100.0%; Score 74; DB 15; Length 261;  
Best Local Similarity 100.0%; Pred. No. 2.8e-05;  
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 TFNTPAMYVAIQAVL 15  
| | | | | | | | | | | | | | | | | | | | | |  
Db 12 TFNTPAMYVAIQAVL 26

RESULT 10  
US-10-425-114-52089  
; Sequence 52089, Application US/10425114  
; Publication No. US20040034888A1  
; GENERAL INFORMATION:  
; APPLICANT: Liu, Jingdong  
; APPLICANT: Zhou, Yihua  
; APPLICANT: Kovalic, David K.  
; APPLICANT: Screen, Steven E  
; APPLICANT: Tabaska, Jack E  
; APPLICANT: Cao, Yongwei  
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated with  
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement  
; FILE REFERENCE: 38-21(53313)B  
; CURRENT APPLICATION NUMBER: US/10/425,114  
; CURRENT FILING DATE: 2003-04-28  
; NUMBER OF SEQ ID NOS: 73128  
; SEQ ID NO 52089  
; LENGTH: 276  
; TYPE: PRT  
; ORGANISM: Glycine max  
; FEATURE:  
; OTHER INFORMATION: Clone ID: 701149157\_FLJ.pep  
US-10-425-114-52089

Query Match 100.0%; Score 74; DB 15; Length 276;  
Best Local Similarity 100.0%; Pred. No. 3e-05;  
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 TFNTPAMYVAIQAVL 15  
| | | | | | | | | | | | | | | | | | | | | |  
Db 27 TFNTPAMYVAIQAVL 41

RESULT 11  
US-10-424-599-278431  
; Sequence 278431, Application US/10424599  
; Publication No. US20040031072A1  
; GENERAL INFORMATION:  
; APPLICANT: La Rosa Thomas J  
; APPLICANT: Kovalic David K  
; APPLICANT: Zhou Yihua  
; APPLICANT: Cao Yongwei  
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated with  
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement  
; FILE REFERENCE: 38-21(53223)B  
; CURRENT APPLICATION NUMBER: US/10/424,599  
; CURRENT FILING DATE: 2003-04-28  
; NUMBER OF SEQ ID NOS: 285684  
; SEQ ID NO 278431  
; LENGTH: 299

; TYPE: PRT  
; ORGANISM: Glycine max  
; FEATURE:  
; OTHER INFORMATION: Clone ID: PAT\_MRT3847\_93446C.1.pep  
US-10-424-599-278431

Query Match 100.0%; Score 74; DB 15; Length 299;  
Best Local Similarity 100.0%; Pred. No. 3.3e-05;  
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 TFNTPAMYVAIQAVL 15  
| | | | | | | | | | | | | | | | | | | | | |  
Db 50 TFNTPAMYVAIQAVL 64

RESULT 12  
US-10-322-281-68  
; Sequence 68, Application US/10322281  
; Publication No. US20040126762A1  
; GENERAL INFORMATION:  
; APPLICANT: David W. Morris  
; APPLICANT: Marc S. Malandro  
; TITLE OF INVENTION: Novel Compositions and Methods in Cancer  
; FILE REFERENCE: 529452001000  
; CURRENT APPLICATION NUMBER: US/10/322,281  
; CURRENT FILING DATE: 2002-12-17  
; NUMBER OF SEQ ID NOS: 866  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 68  
; LENGTH: 371  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-10-322-281-68

Query Match 100.0%; Score 74; DB 16; Length 371;  
Best Local Similarity 100.0%; Pred. No. 4.2e-05;  
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 TENTPAMYVAIQAVL 15  
| | | | | | | | | | | | | | | | | | | | | |  
Db 126 TENTPAMYVAIQAVL 140

RESULT 13  
US-10-205-194-93  
; Sequence 93, Application US/10205194  
; Publication No. US20030134301A1  
; GENERAL INFORMATION:  
; APPLICANT: Warner-Lambert Company  
; APPLICANT: Lee, Kevin  
; APPLICANT: Dixon, Alistair  
; APPLICANT: Brooksbank, Robert  
; APPLICANT: Pincock, Robert  
; TITLE OF INVENTION: Identification and Use of Molecules Implicated in Pain  
; FILE REFERENCE: WL-A-018201  
; CURRENT APPLICATION NUMBER: US/10/205,194  
; CURRENT FILING DATE: 5200-07-24  
; PRIOR APPLICATION NUMBER: GB 0118354.0  
; PRIOR FILING DATE: 2001-07-27  
; NUMBER OF SEQ ID NOS: 177  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 93  
; LENGTH: 375  
; TYPE: PRT  
; ORGANISM: Rattus rattus  
; FEATURE:  
; OTHER INFORMATION: Cytoplasmic gamma isoform of actin  
US-10-205-194-93

Query Match 100.0%; Score 74; DB 14; Length 375;  
Best Local Similarity 100.0%; Pred. No. 4.2e-05;  
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Search completed: April 8, 2005, 12:50:59  
Job time : 42.6667 secs

Qy 1 TFTPAMYVAIQAVL 15  
|||||  
Db 126 TFTPAMYVAIQAVL 140

RESULT 14  
US-10-316-253-88  
; Sequence 88, Application US/10316253  
; Publication No. US20030162706A1  
; GENERAL INFORMATION:  
; APPLICANT: The Procter & Gamble Company  
; APPLICANT: Peters, Kevin  
; APPLICANT: Thompson, Larry  
; APPLICANT: Wang, Feng  
; APPLICANT: Greis, Kenneth  
; TITLE OF INVENTION: Angiogenesis Modulating Proteins  
; FILE REFERENCE: 8865M  
; CURRENT APPLICATION NUMBER: US/10/316,253  
; CURRENT FILING DATE: 2002-12-10  
; PRIOR APPLICATION NUMBER: US 60/355,295  
; PRIOR FILING DATE: 2002-02-08  
; NUMBER OF SEQ ID NOS: 308  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 88  
; LENGTH: 375  
; TYPE: PRT  
; ORGANISM: Rattus norvegicus  
US-10-316-253-88

Query Match 100.0%; Score 74; DB 14; Length 375;  
Best Local Similarity 100.0%; Pred. No. 4.2e-05;  
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 TFTPAMYVAIQAVL 15  
|||||  
Db 126 TFTPAMYVAIQAVL 140

RESULT 15  
US-10-205-331-94  
; Sequence 94, Application US/10205331  
; Publication No. US20040058326A1  
; GENERAL INFORMATION:  
; APPLICANT: Warner-Lambert Company  
; APPLICANT: Lee, Kevin  
; APPLICANT: Dixon, Alistair  
; APPLICANT: Brooksbank, Robert  
; APPLICANT: Pinnoch, Robert  
; TITLE OF INVENTION: Identification and Use of Molecules Implicated in Pain  
; FILE REFERENCE: WL-A-018199  
; CURRENT APPLICATION NUMBER: US/10/205,331  
; CURRENT FILING DATE: 2002-07-24  
; PRIOR APPLICATION NUMBER: GB 0118354.0  
; PRIOR FILING DATE: 2001-07-27  
; NUMBER OF SEQ ID NOS: 117  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 94  
; LENGTH: 375  
; TYPE: PRT  
; ORGANISM: Rattus norvegicus  
; FEATURE:  
; OTHER INFORMATION: Cytoplasmic beta-actin  
US-10-205-331-94

Query Match 100.0%; Score 74; DB 15; Length 375;  
Best Local Similarity 100.0%; Pred. No. 4.2e-05;  
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 TFTPAMYVAIQAVL 15  
|||||  
Db 126 TFTPAMYVAIQAVL 140

**THIS PAGE BLANK (USPTO)**

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: April 8, 2005, 09:58:56 ; Search time 9.33333 Seconds  
(without alignments)  
154.634 Million cell updates/sec

Title: US-09-423-351C-6  
Perfect score: 74  
Sequence: 1 TFNTPAMYVAIQAVL 15

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 283416 seqs, 96216763 residues  
Total number of hits satisfying chosen parameters: 283416

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000  
Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : PIR\_79:\*  
1: pir1.\*  
2: pir2.\*  
3: pir3.\*  
4: pir4.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	74	100.0	195	2 S39777	actin beta - pig (
2	74	100.0	195	2 S20097	actin 85c - potato
3	74	100.0	308	2 A03000	actin 3 - fruit fl
4	74	100.0	327	2 S11452	actin (clone 302)
5	74	100.0	328	2 S05430	actin beta - grass
6	74	100.0	362	2 A26559	actin type 5, cyto
7	74	100.0	374	1 A7BOB	actin beta - bovin
8	74	100.0	374	1 A7BOG	actin gamma - bovi
9	74	100.0	374	2 JC5818	gamma-actin - huma
10	74	100.0	375	1 A7RTC	actin beta - rat
11	74	100.0	375	1 A48324	actin beta, cytosk
12	74	100.0	375	1 A7AX	actin - Acanthamo
13	74	100.0	375	1 A7CHB	actin beta - chick
14	74	100.0	375	1 A7DO	actin - slime mold
15	74	100.0	375	1 A7HUB	actin beta - human
16	74	100.0	375	1 A7HUG	actin gamma 1 - hu
17	74	100.0	375	1 A7MSB	actin beta - mouse
18	74	100.0	375	1 A7MSG	actin gamma - mous
19	74	100.0	375	1 A7RBB	actin beta, non-mu
20	74	100.0	375	1 S11222	actin gamma, cytos
21	74	100.0	375	2 S71125	actin beta-2, cyto
22	74	100.0	375	2 S71124	actin beta-1, cyto
23	74	100.0	375	2 S71126	actin beta, cytosol
24	74	100.0	375	2 A55001	actin beta - goose
25	74	100.0	376	1 A43552	actin gamma, cytos
26	74	100.0	376	1 A7FY	actin - slime mold
27	74	100.0	376	2 A48449	Actin-1A - nematod
28	74	100.0	376	2 JS0189	actin, cytosolic -
29	74	100.0	376	2 C23412	actin 3-sub1 - bli

actin 15 - slime m  
actin 4 - Caenotha  
actin 1 and actin  
actin 2 - Caenortha  
hypothetical prote  
actin - California  
actin (clone 403)  
actin 5 - Atlantic  
actin 11 - Atlanti  
actin 3 - Atlantic  
actin beta, cytosk  
actin A4 - silkwor  
actin A3, cytosoli  
actin 1 - rice  
actin - common tob  
actin 97 - potato

ALIGNMENTS

RESULT 1

S39777  
actin beta - pig (fragments)  
C:Species: Sus scrofa domestica (domestic pig)  
C>Date: 07-Oct-1994 #sequence\_revision 12-Apr-1996 #text\_change 09-Jul-2004  
C:Accession: S39777  
R:Adamski, J.; Husein, B.; Thole, H.H.; Groeschel-Stewart, U.; Jungblut, P.W.  
Biochem. J. 296, 797-802, 1993  
A>Title: Linkage of 17-beta-oestradiol dehydrogenase to actin by epsilon- (gamma-glutamyl  
A:Reference number: S39777; MUID:94107247; PMID:8280079  
A:Accession: S39777  
A:Molecule type: Protein  
A:Residues: 1-20;21-52;53-79;80-85;86-95;96-116;117-151;152-159;160-195 <ADA>  
A:Cross-references: UNIPROT:Q7M3B0  
C:Superfamily: actin

Query Match 100.0%; Score 74; DB 2; Length 195;  
Best Local Similarity 100.0%; Pred. No. 5.2e-06;  
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 TFNTPAMYVAIQAVL 15  
Db 60 TFNTPAMYVAIQAVL 74

RESULT 2

S20097  
actin 85c - potato (fragment)  
C:Species: Solanum tuberosum (potato)  
C>Date: 22-Nov-1993 #sequence\_revision 17-Oct-1997 #text\_change 09-Jul-2004  
C:Accession: S20097  
R:Drouin, G.; Dover, G.A.  
J. Mol. Evol. 31, 132-150, 1990  
A>Title: Independent gene evolution in the potato actin gene family demonstrated by phyl  
A:Reference number: S20092; MUID:91012599; PMID:2120451  
A:Accession: S20097  
A:Molecule type: DNA  
A:Residues: 1-195 <DRO>  
A:Cross-references: UNIPROT:P30170; EMBL:X55747; NID:g21541; PIDN:CAA39277.1; PID:g13455  
C:Genetics:  
A:Introns: 132/1  
C:Superfamily: actin  
C:Keywords: cytoskeleton; structural protein

Query Match 100.0%; Score 74; DB 2; Length 195;  
Best Local Similarity 100.0%; Pred. No. 5.2e-06;  
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 TFNTPAMYVAIQAVL 15  
Db 108 TFNTPAMYVAIQAVL 122

```
RESULT 3
A03000
actin 3 - fruit fly (Drosophila melanogaster) (fragments)
C;Species: Drosophila melanogaster
C;Date: 31-Mar-1991 #sequence_revision 31-Mar-1991 #text_change 09-Jul-2004
C;Accession: A03000
R;Fyrberg, E.A.; Bond, B.J.; Hershey, N.D.; Mixter, K.S.; Davidson, N.
Cell 24, 107-116, 1981
A;Title: The actin genes of Drosophila: protein coding regions are highly conserved but
A;Reference number: A03000; MUID:81210174; PMID:6263481
A;Accession: A03000
A;Molecule type: DNA
A;Residues: 1-308 <FYR>
A;Cross-references: UNIPROT:P02572
A;Note: there are 68 unsequenced residues between positions 160 and 161. Partial sequence
C;Genetics:
A;Gene: FlyBase:Act42A
A;Cross-references: FlyBase:FBgn0000043
A;Map position: 42A
C;Superfamily: actin
C;Keywords: methylated amino acid
F;74/Modified site: 3'-methylhistidine (His) #status predicted

Query Match 100.0%; Score 74; DB 2; Length 308;
Best Local Similarity 100.0%; Pred. No. 8.6e-06;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 TFNTPAMYVAIQAVL 15
Db 127 TFNTPAMYVAIQAVL 141

RESULT 4
S11452
actin (clone 302) - brine shrimp (fragment)
C;Species: Artemia sp. (brine shrimp)
C;Date: 17-Apr-1993 #sequence_revision 17-Apr-1993 #text_change 09-Jul-2004
C;Accession: S11452
R;Macias, M.T.; Sastre, L.
Nucleic Acids Res. 18, 5219-5225, 1990
A;Title: Molecular cloning and expression of four actin isoforms during Artemia developm
A;Reference number: S11450; MUID:90384823; PMID:2402445
A;Accession: S11452
A;Molecule type: mRNA
A;Residues: 1-327 <MAC>
A;Cross-references: UNIPROT:P18602; EMBL:X52604; NID:G5666; PIDN:CAA36837.1; PID:g829162
C;Superfamily: actin
C;Keywords: methylated amino acid; structural protein
F;25/Modified site: 3'-methylhistidine (His) #status predicted

Query Match 100.0%; Score 74; DB 2; Length 327;
Best Local Similarity 100.0%; Pred. No. 9.2e-06;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 TFNTPAMYVAIQAVL 15
Db 78 TFNTPAMYVAIQAVL 92

RESULT 5
S05430
actin beta - grass carp
C;Species: Ctenopharyngodon idella (grass carp)
C;Date: 07-Jun-1990 #sequence_revision 07-Jun-1990 #text_change 09-Jul-2004
C;Accession: S05430
R;Liu, Z.; Zhu, Z.; Roberg, K.; Faras, A.J.; Guise, K.S.; Kapuscinski, A.R.; Hackett, P.
Nucleic Acids Res. 17, 5850, 1989
A;Title: The beta-actin gene of carp (Ctenopharyngodon idella).
A;Reference number: S05430; MUID:89345185; PMID:2762162
A;Accession: S05430
A;Status: translation not shown
```

```
A;Molecule type: DNA
A;Residues: 1-328 <LIU>
A;Cross-references: UNIPROT:P83751; EMBL:M25013
C;Genetics:
A;Introns: 41/3; 121/3; 268/1
C;Superfamily: actin
C;Keywords: cytoskeleton; methylated amino acid
F;73/Modified site: 3'-methylhistidine (His) #status predicted

Query Match 100.0%; Score 74; DB 2; Length 328;
Best Local Similarity 100.0%; Pred. No. 9.2e-06;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 TENTPAMYVAIQAVL 15
Db 126 TFNTPAMYVAIQAVL 140

RESULT 6
A26559
actin type 5, cytosolic - chicken
C;Species: Gallus gallus (chicken)
C;Date: 09-Sep-1987 #sequence_revision 09-Sep-1987 #text_change 05-Dec-1997
C;Accession: A26559
R;Bergsma, D.J.; Chang, K.S.; Schwartz, R.J.
Mol. Cell. Biol. 5, 1151-1162, 1985
A;Reference number: A26559; MUID:85213487; PMID:4000121
A;Accession: A26559
A;Molecule type: DNA
A;Residues: 1-362 <BER>
C;Superfamily: actin
C;Keywords: cytosol; methylated amino acid
F;74/Modified site: 3'-methylhistidine (His) #status predicted

Query Match 100.0%; Score 74; DB 2; Length 362;
Best Local Similarity 100.0%; Pred. No. 1e-05;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 TENTPAMYVAIQAVL 15
Db 127 TFNTPAMYVAIQAVL 141

RESULT 7
ATB08
actin beta - bovine (tentative sequence)
C;Species: Bos primigenius taurus (cattle)
C;Date: 30-Sep-1991 #sequence_revision 30-Sep-1991 #text_change 09-Jul-2004
C;Accession: E14185; A39105; A02999; A14185
R;Vandekerckhove, J.; Weber, K.
Eur. J. Biochem. 90, 451-462, 1978
A;Title: Actin amino-acid sequences. Comparison of actins from calf thymus, bovine brain
A;Reference number: A14185; MUID:79045349; PMID:213279
A;Accession: E14185
A;Molecule type: protein
A;Residues: 1-374 <VAN>
A;Cross-references: UNIPROT:P60712
A;Note: only peptides that differed in composition from the corresponding peptides of rai
R;Degen, J.L.; Neubauer, M.G.; Degen, S.U.F.; Seyfried, C.E.; Morris, D.R.
J. Biol. Chem. 258, 12153-12162, 1983
A;Title: Regulation of protein synthesis in mitogen-activated bovine lymphocytes. Analys
A;Reference number: A39105; MUID:84032385; PMID:6195151
A;Accession: A39105
A;Molecule type: mRNA
A;Residues: 76-227;344-374 <DEG>
A;Cross-references: GB:K00622; GB:K00623
A;Note: actins beta and gamma were not distinguished in this study
C;Comment: Vertebrate nonmuscle cells contain two highly conserved cytoskeletal actins,
C;Superfamily: actin
C;Keywords: blocked amino end; cell motility; cytoskeleton; methylated amino acid; micro
F;1/Modified site: blocked amino end (Asp) (probably acetylated) #status experimental
F;72/Modified site: 3'-methylhistidine (His) #status predicted
```



Query Match 100.0%; Score 74; DB 1; Length 374;  
 Best Local Similarity 100.0%; Pred. No. 1.1e-05;  
 Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 TFNTPAMYVAIQAVL 15  
 |||||  
 Db 125 TFNTPAMYVAIQAVL 139

RESULT 8  
 ATBOG  
 actin gamma - bovine (tentative sequence)  
 C:Species: Bos primigenius taurus (cattle)  
 C:Date: 30-Sep-1991 #sequence\_revision 30-Sep-1991 #text\_change 09-Jul-2004  
 C:Accession: B14185; A02999  
 R:Vanderkerckhove, J.; Weber, K.  
 Eur. J. Biochem. 90, 451-462, 1978  
 A:Title: Actin amino-acid sequences. Comparison of actins from calf thymus, bovine brain  
 A:Reference number: A14185; MUID:79045349; PMID:213279  
 A:Accession: B14185  
 A:Molecule type: protein  
 A:Residues: 1-374 <VAN>  
 A:Cross-references: UNIPROT:P02571  
 A:Note: only peptides that differed in composition from the corresponding peptides of rat  
 C:Comment: Vertebrate nonmuscle cells contain two highly conserved cytoskeletal actins,  
 C:Superfamily: actin  
 C:Keywords: blocked amino end; cell motility; cytoskeleton; methylated amino acid; micro  
 F:1/Modified site: blocked amino end (Glu) (probably acetylated) #status experimental  
 F:72/Modified site: 3'-methylhistidine (His) #status predicted

Query Match 100.0%; Score 74; DB 1; Length 374;  
 Best Local Similarity 100.0%; Pred. No. 1.1e-05;  
 Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 TFNTPAMYVAIQAVL 15  
 |||||  
 Db 125 TFNTPAMYVAIQAVL 139

RESULT 9  
 JC5818  
 gamma-actin - human  
 C:Species: Homo sapiens (man)  
 C:Date: 20-Feb-1998 #sequence\_revision 13-Mar-1998 #text\_change 09-Jul-2004  
 C:Accession: JC5818; PC4501  
 R:Hauschildt, S.; Schwarz, C.; Heine, H.; Ulmer, A.J.; Flad, H.D.; Rietschel, E.T.; Jene  
 Biochem. Biophys. Res. Commun. 241, 670-674, 1997  
 A:Title: Actin: A target of lipopolysaccharid-induced phosphorylation in human monocytes  
 A:Reference number: JC5818; MUID:98096379; PMID:9434766  
 A:Accession: JC5818  
 A:Molecule type: protein  
 A:Residues: 1-374 <HAU>  
 A:Cross-references: UNIPROT:P02571  
 A:Experimental source: monocyte  
 A:Accession: PC4501  
 A:Molecule type: protein  
 A:Residues: 1-61;84-112;147-190;196-209;215-253;335-358 <HA2>  
 A:Experimental source: monocyte  
 C:Comment: This protein is involved in a signal transduction that eventually leads to m  
 C:Superfamily: actin

Query Match 100.0%; Score 74; DB 2; Length 374;  
 Best Local Similarity 100.0%; Pred. No. 1.1e-05;  
 Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 TFNTPAMYVAIQAVL 15  
 |||||  
 Db 125 TFNTPAMYVAIQAVL 139

RESULT 10  
 ATRTC  
 actin beta - rat

C:Species: Rattus norvegicus (Norway rat)  
 C:Date: 13-Jun-1983 #sequence\_revision 13-Jun-1983 #text\_change 22-Jun-1999  
 C:Accession: A38571; A02999  
 R:Nudel, U.; Zakut, R.; Shani, M.; Neuman, S.; Levy, Z.; Yaffe, D.  
 Nucleic Acids Res. 11, 1759-1771, 1983  
 A:Title: The nucleotide sequence of the rat cytoplasmic beta-actin gene.  
 A:Reference number: A38571; MUID:83168920; PMID:6300777  
 A:Accession: A38571  
 A:Molecule type: DNA  
 A:Residues: 1-375 <NUD>  
 A:Cross-references: GB:J00691; NID:G202653; PIDN:AAA40657.1; PID:G202654  
 C:Comment: Vertebrate nonmuscle cells contain two highly conserved cytoskeletal actins,  
 C:Genetics:  
 A:Introns: 41/3; 121/3; 268/1; 328/3  
 C:Superfamily: actin  
 C:Keywords: cell motility; cytoskeleton; methylated amino acid; microfilament; mitosis;  
 F:2-375/Product: actin beta #status predicted <WAT>  
 F:73/Modified site: 3'-methylhistidine (His) #status predicted

Query Match 100.0%; Score 74; DB 1; Length 375;  
 Best Local Similarity 100.0%; Pred. No. 1.1e-05;  
 Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 TFNTPAMYVAIQAVL 15  
 |||||  
 Db 126 TFNTPAMYVAIQAVL 140

RESULT 11  
 A48324  
 actin beta, cytoskeletal - common carp  
 C:Species: Cyprinus carpio (common carp)  
 C:Date: 03-Feb-1994 #sequence\_revision 11-Apr-1997 #text\_change 09-Jul-2004  
 C:Accession: A48324  
 R:Liu, Z.; Zhu, Z.; Roberg, K.; Faras, A.; Guise, K.; Kapuscinski, A.R.; Hackett, P.B.  
 DNA Seq. 1, 125-136, 1990  
 A:Title: Isolation and characterization of beta-actin gene of carp (Cyprinus carpio).  
 A:Reference number: A48324; MUID:92190540; PMID:2134183  
 A:Accession: A48324  
 A:Molecule type: DNA  
 A:Residues: 1-375 <LIU>  
 A:Cross-references: UNIPROT:P83750; GB:M24113; NID:G213041; PIDN:AAA68896.1; PID:G213042  
 A:Note: the authors translated the codon TTC for residue 21 as Pro, AAG for residue 50 a  
 7 as Pro  
 A:Note: the authors failed to translated the codon GGT for residue 42 as Gly  
 C:Comment: Vertebrate nonmuscle cells contain two highly conserved cytoskeletal actins,  
 C:Genetics:  
 A:Introns: 41/3; 121/3; 268/3; 328/3  
 C:Superfamily: actin  
 C:Keywords: acetylated amino end; cell motility; cytoskeleton; methylated amino acid; mi  
 F:2-375/Product: actin beta, cytoskeletal #status predicted <WAT>  
 F:2/Modified site: acetylated amino end (Asp) (in mature form) #status predicted  
 F:73/Modified site: 3'-methylhistidine (His) #status predicted

Query Match 100.0%; Score 74; DB 1; Length 375;  
 Best Local Similarity 100.0%; Pred. No. 1.1e-05;  
 Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 TFNTPAMYVAIQAVL 15  
 |||||  
 Db 126 TFNTPAMYVAIQAVL 140

RESULT 12  
 ATAX  
 actin - Acanthamoeba castellanii  
 C:Species: Acanthamoeba castellanii  
 C:Date: 31-Dec-1991 #sequence\_revision 31-Dec-1991 #text\_change 09-Jul-2004  
 C:Accession: A92886; A03004  
 R:Nellen, W.; Gallwitz, D.  
 J. Mol. Biol. 159, 1-18, 1982  
 A:Title: Actin genes and actin messenger RNA in Acanthamoeba castellanii. Nucleotide se  
 A:Reference number: A92886; MUID:83033627; PMID:6290670

A;Accession: A92886  
A;Molecule type: DNA  
A;Residues: 1-375 <NLE>  
A;Cross-references: UNIPROT:P02578; GB:V00002; GB:J01016; NID:G5565; PIDN:CAA23399.1; PI  
C;Comment: There are at least three actin genes in A. castellanii.  
C;Genetics:  
A;Introns: 105/3  
C;Superfamily: actin  
C;Keywords: methylated amino acid  
F;73/Modified site: 3'-methylhistidine (His) #status predicted

Query Match 100.0%; Score 74; DB 1; Length 375;  
Best Local Similarity 100.0%; Pred. No. 1.1e-05;  
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 TTTNPAMYVAIQAVL 15  
|||||  
Db 126 TTTNPAMYVAIQAVL 140

RESULT 13  
ATHUB  
actin beta - chicken  
C;Species: Gallus gallus (chicken)  
C;Date: 31-Mar-1993 #sequence\_revision 31-Mar-1993 #text\_change 09-Jul-2004  
C;Accession: A20888; I50154  
R;Kost, T.A.; Theodorakis, N.; Hughes, S.H.  
Nucleic Acids Res. 11, 8287-8301, 1983  
A;Title: The nucleotide sequence of the chick cytoplasmic beta-actin gene.  
A;Reference number: A20888; MUID:84169478; PMID:6324080  
A;Accession: A20888  
A;Molecule type: DNA  
A;Residues: 1-375 <KOS>  
A;Cross-references: UNIPROT:P60706; EMBL:X00182  
A;Note: the sequence shown follows the authors' translation at position 336  
R;Chang, K.  
Mol. Cell. Biol. 4, 2498-2508, 1984  
A;Title: Isolation and characterization of six different chicken actin genes.  
A;Reference number: I50153; MUID:85085956; PMID:6513927  
A;Accession: I50154  
A;Status: translated from GB/EMBL/DBJ  
A;Molecule type: DNA  
A;Residues: 1-17 <CHA>  
A;Cross-references: GB:K02259; NID:G211086; PIDN:AAA49572.1; PID:G211087  
C;Comment: Vertebrate nonmuscle cells contain two highly conserved cytoskeletal actins,  
C;Genetics:  
A;Introns: 41/3; 121/3; 268/1; 328/3  
C;Superfamily: actin  
C;Keywords: acetylated amino end; cell motility; cytoskeleton; methylated amino acid; mi  
F;2-375/Product: actin beta #status predicted <WAT>  
F;2/Modified site: acetylated amino end (Asp) (in mature form) #status predicted  
F;73/Modified site: 3'-methylhistidine (His) #status predicted

Query Match 100.0%; Score 74; DB 1; Length 375;  
Best Local Similarity 100.0%; Pred. No. 1.1e-05;  
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 TTTNPAMYVAIQAVL 15  
|||||  
Db 126 TTTNPAMYVAIQAVL 140

RESULT 14  
AIDO  
actin - slime mold (Dictyostelium discoideum)  
C;Species: Dictyostelium discoideum  
C;Date: 31-Dec-1991 #sequence\_revision 31-Dec-1991 #text\_change 09-Jul-2004  
C;Accession: A93223; A92871; S67999; A03004  
R;Vandekerckhove, J.; Weber, K.  
Nature 284, 475-477, 1980  
A;Title: Vegetative Dictyostelium cells containing 17 actin genes express a single major  
A;Reference number: A93223; MUID:80143270; PMID:6892652  
A;Accession: A93223

A;Molecule type: protein  
A;Residues: 1-375 <VAN>  
A;Cross-references: UNIPROT:P02577  
R;McKeown, M.; Firtel, R.A.  
J. Mol. Biol. 151, 593-606, 1981  
A;Title: Evidence for sub-families of actin genes in Dictyostelium as determined by comp  
A;Reference number: A92871; MUID:82122583; PMID:6276562  
A;Accession: A92871  
A;Molecule type: mRNA; DNA  
A;Residues: 288-375 <WCK>  
R;Jungbluth, A.; Eckerskorn, C.; Gerisch, G.; Lottspeich, F.; Stocker, S.; Schweiger, A.  
FEBS Lett. 375, 87-90, 1995  
A;Title: Stress-induced tyrosine phosphorylation of actin in Dictyostelium cells and loc  
A;Reference number: S67999; MUID:96087090; PMID:7498488  
A;Accession: S67999  
A;Molecule type: protein  
A;Residues: 51-61 <JUN>  
C;Comment: Although Dictyostelium may contain 17 actin genes, only one major actin is ex;  
C;Superfamily: actin  
C;Keywords: methylated amino acid  
F;73/Modified site: 3'-methylhistidine (His) #status predicted

Query Match 100.0%; Score 74; DB 1; Length 375;  
Best Local Similarity 100.0%; Pred. No. 1.1e-05;  
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 TTTNPAMYVAIQAVL 15  
|||||  
Db 126 TTTNPAMYVAIQAVL 140

RESULT 15  
ATHUB  
actin beta - human  
C;Species: Homo sapiens (man)  
C;Date: 30-Sep-1991 #sequence\_revision 30-Sep-1991 #text\_change 09-Jul-2004  
C;Accession: A25168; A37248; S23707; I39394; S38782; A02999  
R;Nakajima-Iijima, S.; Hamada, H.; Reddy, P.; Kakunaga, T.  
Proc. Natl. Acad. Sci. U.S.A. 82, 6133-6137, 1985  
A;Title: Molecular structure of the human cytoplasmic beta-actin gene: interspecies homo  
A;Reference number: A25168; MUID:85298307; PMID:2994062  
A;Accession: A25168  
A;Molecule type: DNA  
A;Residues: 1-375 <NKA>  
A;Cross-references: UNIPROT:P60709; GB:M10277; NID:G177967; PIDN:AAA51567.1; PID:G177968  
A;Note: the authors translated the codon CAG for residue 137 as Glu  
R;Ng, S.Y.; Gunning, P.; Eddy, R.; Ponte, P.; Leavitt, J.; Shows, T.; Keddes, L.  
Mol. Cell. Biol. 5, 2720-2732, 1985  
A;Title: Evolution of the functional human beta-actin gene and its multi-pseudogene fami  
A;Reference number: A37248; MUID:86284634; PMID:3837182  
A;Accession: A37248  
A;Molecule type: mRNA  
A;Residues: 1-375 <NGX>  
A;Cross-references: EMBL:X00351; DBJ:J00074; GB:M10278; NID:G28251; PIDN:CAA25099.1; PI  
A;Note: The human genome contains one functional beta-actin gene and a number of process;  
R;Ponte, P.; Ng, S.Y.; Engel, J.; Gunning, P.; Keddes, L.  
Nucleic Acids Res. 12, 1687-1696, 1984  
A;Title: Evolutionary conservation in the untranslated regions of actin mRNAs: DNA sequ  
A;Reference number: S23707; MUID:84144061; PMID:6322116  
A;Accession: S23707  
A;Status: preliminary  
A;Molecule type: mRNA  
A;Residues: 1-375 <PON>  
A;Cross-references: EMBL:X63432; NID:G28335; PIDN:CAA45026.1; PID:G28336  
R;Hanukoglu, I.; Tanese, N.; Fuchs, E.  
J. Mol. Biol. 163, 673-678, 1983  
A;Title: Complementary DNA sequence of a human cytoplasmic actin. Interspecies divergenc.  
A;Reference number: A37247; MUID:83189093; PMID:6842590  
A;Accession: A37247  
A;Molecule type: mRNA  
A;Residues: 252-375 <HAN>  
A;Cross-references: EMBL:V00478; NID:G28244; PIDN:CAA23745.1; PID:G825616  
R;Gunning, P.; Ponte, P.; Okayama, H.; Engel, J.; Blau, H.; Keddes, L.

Mol. Cell. Biol. 3, 787-795, 1983  
A:Title: Isolation and characterization of full-length cDNA clones for human alpha-, beta- and gamma-actin  
A:Reference number: I39394; MUID:83244575; PMID:6865942  
A:Accession: I39394  
A:Status: preliminary; translated from GB/EMBL/DBJ  
A:Molecule type: mRNA  
A:Residues: 1-17 <GUN>  
A:Cross-references: GB:K00790; NID:gl78031; PIDN:AAA51578.1; PID:gl78032  
R:Ohmuri, H.  
submitted to the EMBL Data Library, October 1991  
A:Reference number: S38782  
A:Accession: S38782  
A:Status: preliminary  
A:Molecule type: mRNA  
A:Residues: 1-138, 'M', 140-294, 'D', 296-375 <OHM>  
A:Cross-references: EMBL:X63432; NID:g28335; PIDN:CAA5026.1; PID:g28336  
C:Comment: Vertebrate nonmuscle cells contain two highly conserved cytoskeletal actins,  
C:Genetics:  
A:Gene: GDB:ACTB  
A:Cross-references: GDB:118964; OMIM:102630  
A:Map position: 7p22-7p22  
A:Introns: 41/3; 121/3; 268/1; 328/3  
C:Superfamily: actin  
C:Keywords: cell motility; cytoskeleton; methylated amino acid; microfilament; mitosis;  
F:2-375/Product: actin beta #status predicted <MAT>  
F:73/Modified site: 3'-methylhistidine (His) #status predicted

Query Match 100.0%; Score 74; DB 1; Length 375;  
Best Local Similarity 100.0%; Pred. No. 1.1e-05;  
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 TTNTPMYVIAQAVL 15  
|||||  
Db 126 TTNTPMYVIAQAVL 140

Search completed: April 8, 2005, 10:53:05  
Job time : 9.3333 secs

**THIS PAGE BLANK (USPTO)**

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: April 8, 2005, 09:58:56 ; Search time 57.8667 Seconds  
(without alignments)  
132.739 Million cell updates/sec

Title: US-09-423-351C-6  
Perfect score: 74  
Sequence: 1 TFNTPAMVAIQAVL 15

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 1612378 seqs, 512079187 residues

Total number of hits satisfying chosen parameters: 1612378

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000  
Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : UniProt\_03:\*  
1: uniprot\_sprot:\*  
2: uniprot\_trembl:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	74	100.0	58	2 Q28242	Q28242 cervus elap
2	74	100.0	60	2 Q18B1	Q18B1 tubifex tub
3	74	100.0	60	2 Q18C0	Q18C0 thelohanell
4	74	100.0	60	2 Q18C2	Q18C2 myxobolus e
5	74	100.0	60	2 Q18C3	Q18C3 myxobolus e
6	74	100.0	60	2 Q18C4	Q18C4 myxobolus m
7	74	100.0	60	2 Q18C5	Q18C5 myxobolus s
8	74	100.0	60	2 Q18C6	Q18C6 myxobolus i
9	74	100.0	68	2 Q18C7	Q18C7 sparus aura
10	74	100.0	95	2 Q18R2	Q18R2 spodoptera
11	74	100.0	97	2 Q9JLX9	Q9JLX9 rattus norv
12	74	100.0	109	2 Q28916	Q28916 macaca fusc
13	74	100.0	117	2 Q8T6P7	Q8T6P7 elysia chlo
14	74	100.0	117	2 Q80012	Q80012 hyla japoni
15	74	100.0	117	2 Q80013	Q80013 hyla japoni
16	74	100.0	122	2 Q94IA3	Q94IA3 phaseolus v
17	74	100.0	125	2 Q802E1	Q802E1 zoarces viv
18	74	100.0	125	2 Q802E2	Q802E2 pachycara b
19	74	100.0	130	2 Q865G0	Q865G0 capra hircu
20	74	100.0	130	2 Q9DE58	Q9DE58 coryphaenoi
21	74	100.0	132	2 Q9IA84	Q9IA84 stizostedio
22	74	100.0	133	2 Q9U4L7	Q9U4L7 agestaspis
23	74	100.0	133	2 Q9U4L9	Q9U4L9 agestaspis
24	74	100.0	136	2 Q95V64	Q95V64 ixodes scap
25	74	100.0	144	2 Q90YX9	Q90YX9 fundulus he
26	74	100.0	145	2 Q91L42	Q91L42 trebouxia j
27	74	100.0	147	2 Q7X9B5	Q7X9B5 fragaria an
28	74	100.0	150	2 Q68AX1	Q68AX1 cynops pyrr
29	74	100.0	157	2 Q964X3	Q964X3 cynops max
30	74	100.0	160	2 Q8T6A7	Q8T6A7 trichinella
31	74	100.0	160	2 Q9NAS3	Q9NAS3 blattella g

32	74	100.0	162	2 Q6UBA0	Q6UBA0 chelonina my
33	74	100.0	165	2 Q8QFS7	Q8QFS7 scylliorhinu
34	74	100.0	170	2 Q9NB01	Q9NB01 metridium s
35	74	100.0	171	2 Q7XAK3	Q7XAK3 asparagus o
36	74	100.0	171	2 Q8JH59	Q8JH59 chelydra se
37	74	100.0	173	2 Q8RX10	Q8RX10 ficus caric
38	74	100.0	174	2 Q819W1	Q819W1 biomphalari
39	74	100.0	174	2 Q819W2	Q819W2 biomphalari
40	74	100.0	174	2 Q819W3	Q819W3 biomphalari
41	74	100.0	174	2 Q819W4	Q819W4 biomphalari
42	74	100.0	174	2 Q819W5	Q819W5 biomphalari
43	74	100.0	174	2 Q819W6	Q819W6 helisoma an
44	74	100.0	174	2 Q819W7	Q819W7 helisoma tr
45	74	100.0	174	2 Q819W8	Q819W8 biomphalari

## ALIGNMENTS

RESULT 1  
Q28242 PRELIMINARY; PRT; 58 AA.  
AC Q28242;  
DT 01-NOV-1996 (TrEMBLrel. 01, Created)  
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)  
DE 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)  
DE Actin, cytoplasmic 1 (Beta-actin) (Fragment).  
GN Name=ACTB;  
OS Cervus elaphus (Red deer).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Cervidae;  
OC Cervinae; Cervus.  
OX NCBI\_TaxID=9860;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC TISSUE=Antler;  
RX MEDLINE=98233260; PubMed=9571767;  
RX DOI=10.1002/(SICI)1097-010X(19980501)281:1<36::AID-JE26>3.0.CO;2-D;  
RA Francis S.M., Suttie J.M.;  
RT "Detection of growth factors and proto-oncogene mRNA in the growing tip of red deer (Cervus elaphus) antler using reverse-transcriptase polymerase chain reaction (RT-PCR).";  
RT J. Exp. Zool. 281:36-42(1998).  
RL -!- FUNCTION: Actins are highly conserved proteins that are involved in various types of cell motility and are ubiquitously expressed in all eukaryotic cells.  
CC -!- SUBUNIT: Polymerization of globular actin (G-actin) leads to a structural filament (F-actin) in the form of a two-stranded helix. Each actin can bind to 4 others.  
CC -!- SUBCELLULAR LOCATION: Cytoplasmic.  
CC -!- MISCELLANEOUS: In vertebrates 3 main groups of actin isoforms, alpha, beta and gamma have been identified. The alpha actins are found in muscle tissues and are a major constituent of the contractile apparatus. The beta and gamma actins co-exist in most cell types as components of the cytoskeleton and as mediators of internal cell motility.  
CC -!- SIMILARITY: Belongs to the actin family.  
DR EMBL; U62112; AB05258.1; -.  
DR HSSP; P02577; INMI.  
DR GO; GO:0015629; C:actin cytoskeleton; IEA.  
DR GO; GO:0005200; F:structural constituent of cytoskeleton; IEA.  
DR InterPro; IPR004001; Actin.  
DR InterPro; IPR004000; Actin\_like.  
DR Pfam; PF00022; Actin; 1.  
DR PROSITE; PS00406; ACTINS\_1; PARTIAL.  
DR PROSITE; PS00432; ACTINS\_2; PARTIAL.  
DR PROSITE; PS01132; ACTINS\_ACT LIKE; PARTIAL.  
KW Cytoskeleton; Multigene family; Structural protein.  
FT NON TER 1 1  
FT CHAIN <1 >58 ACTIN, CYTOPLASMIC 1.  
FT NON TER 58 58  
SQ SEQUENCE 58 AA; 6290 MW; 26D0D8BEC629A61 CRC64;

```
Query Match      100.0%; Score 74; DB 2; Length 59;
Best Local Similarity 100.0%; Pred. No. 5.5e-06;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 TFNTPMYVAIQAVL 15
Db 7 TFNTPMYVAIQAVL 21

RESULT 2
Q818B1 ID Q818B1 PRELIMINARY; PRT; 60 AA.
AC Q818B1;
DT 01-MAR-2003 (TrEMBLrel. 23, Created)
DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Putative beta-actin (Fragment).
OS Tubifex tubifex (Sludge worm).
OC Eukaryota; Metazoa; Annelida; Clitellata; Oligochaeta; Haplotaxida;
OC Tubificina; Tubificidae; Tubifex.
OX NCBI_TaxID=6396;
RN [1]
RP SEQUENCE FROM N.A.
RA Kelley G.O., Beauchamp K.A., Hedrick R.R.;
RL Submitted (SEP-2002) to the EMBL/GenBank/DBJ databases.
CC -!- FUNCTION: Actins are highly conserved proteins that are involved
CC in various types of cell motility and are ubiquitously expressed
CC in all eukaryotic cells (By similarity).
CC -!- SIMILARITY: Belongs to the actin family.
DR EMBL; AY157023; AAN85108.1; -.
DR HSSP; P02568; 1MA9.
DR GO; GO:0015629; C:actin cytoskeleton; IEA.
DR GO; GO:0005200; F:structural constituent of cytoskeleton; IEA.
DR InterPro; IPR004000; Actin_like.
DR Pfam; PF00022; Actin; 1.
DR PRINTS; PR00190; ACTIN.
DR PROSITE; PS01132; ACTINS_ACT_LIKE; 1.
KW Structural protein.
FT NON_TER 1 60
FT NON_TER 60 60
SQ SEQUENCE 60 AA; 6591 MW; 8127546708884838 CRC64;

Query Match      100.0%; Score 74; DB 2; Length 60;
Best Local Similarity 100.0%; Pred. No. 5.5e-06;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 TFNTPMYVAIQAVL 15
Db 26 TFNTPMYVAIQAVL 40

RESULT 3
Q818C0 ID Q818C0 PRELIMINARY; PRT; 60 AA.
AC Q818C0;
DT 01-MAR-2003 (TrEMBLrel. 23, Created)
DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Putative beta-actin (Fragment).
OS Theohanelus nikolskii.
OC Eukaryota; Metazoa; Myxozoa; Myxosporia; Bivalvulida; Platysporina;
OC Myxobolidae; Theohanelidae.
OX NCBI_TaxID=215725;
RN [1]
RP SEQUENCE FROM N.A.
RA Kelley G.O., Beauchamp K.A., Hedrick R.R.;
RL Submitted (SEP-2002) to the EMBL/GenBank/DBJ databases.
CC -!- FUNCTION: Actins are highly conserved proteins that are involved
CC in various types of cell motility and are ubiquitously expressed
CC in all eukaryotic cells (By similarity).
CC -!- SIMILARITY: Belongs to the actin family.
DR EMBL; AY157014; AAN85099.1; -.
DR HSSP; P02577; 1NN1.

Query Match      100.0%; Score 74; DB 2; Length 60;
Best Local Similarity 100.0%; Pred. No. 5.5e-06;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 TFNTPMYVAIQAVL 15
Db 26 TFNTPMYVAIQAVL 40

RESULT 4
Q818C2 ID Q818C2 PRELIMINARY; PRT; 60 AA.
AC Q818C2;
DT 01-MAR-2003 (TrEMBLrel. 23, Created)
DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Putative beta-actin (Fragment).
OS Myxobolus episcumalis.
OC Eukaryota; Metazoa; Myxozoa; Myxosporia; Bivalvulida; Platysporina;
OC Myxobolidae; Myxobolus.
OX NCBI_TaxID=204748;
RN [1]
RP SEQUENCE FROM N.A.
RA Kelley G.O., Beauchamp K.A., Hedrick R.R.;
RL Submitted (SEP-2002) to the EMBL/GenBank/DBJ databases.
CC -!- FUNCTION: Actins are highly conserved proteins that are involved
CC in various types of cell motility and are ubiquitously expressed
CC in all eukaryotic cells (By similarity).
CC -!- SIMILARITY: Belongs to the actin family.
DR EMBL; AY157012; AAN85097.1; -.
DR HSSP; P02577; 1NN1.
DR GO; GO:0015629; C:actin cytoskeleton; IEA.
DR GO; GO:0005200; F:structural constituent of cytoskeleton; IEA.
DR InterPro; IPR004000; Actin_like.
DR Pfam; PF00022; Actin; 1.
DR PRINTS; PR00190; ACTIN.
DR PROSITE; PS01132; ACTINS_ACT_LIKE; 1.
KW Structural protein.
FT NON_TER 1 60
FT NON_TER 60 60
SQ SEQUENCE 60 AA; 6484 MW; 3D93D5A1D3D424A7 CRC64;

Query Match      100.0%; Score 74; DB 2; Length 60;
Best Local Similarity 100.0%; Pred. No. 5.5e-06;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 TFNTPMYVAIQAVL 15
Db 26 TFNTPMYVAIQAVL 40

RESULT 5
Q818C3 ID Q818C3 PRELIMINARY; PRT; 60 AA.
AC Q818C3;
DT 01-MAR-2003 (TrEMBLrel. 23, Created)
DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Putative beta-actin (Fragment).
OS Myxobolus exiguus.
OC Eukaryota; Metazoa; Myxozoa; Myxosporia; Bivalvulida; Platysporina;
OC Myxobolidae; Myxobolus.
```

```
OX NCBI_TaxID=204750;
RN [1]
RP SEQUENCE FROM N.A.
RA Kelley G.O., Beauchamp K.A., Hedrick R.R.;
RL Submitted (SEP-2002) to the EMBL/GenBank/DBJ databases.
CC -!- FUNCTION: Actins are highly conserved proteins that are involved
CC in various types of cell motility and are ubiquitously expressed
CC in all eukaryotic cells (By similarity).
CC -!- SIMILARITY: Belongs to the actin family.
DR EMBL; AY157011; AAN85096.1; -.
DR HSSP; P02577; 1NM1.
DR GO; GO:0015629; C:actin cytoskeleton; IEA.
DR InterPro; IPR004000; Actin_like.
DR PRINTS; PR00190; ACTIN.
KW Structural protein.
FT NON_TER 1 60
FT NON_TER 60 60
SQ SEQUENCE 60 AA; 6532 MW; C383DFABD3D424AA CRC64;

Query Match 100.0%; Score 74; DB 2; Length 60;
Best Local Similarity 100.0%; Pred. No. 5.5e-06;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 TFNTPMYVAIQAVL 15
Db 26 TFNTPMYVAIQAVL 40

RESULT 6
Q818C4
ID Q818C4 PRELIMINARY; PRT; 60 AA.
AC Q818C4;
DT 01-MAR-2003 (TrEMBLrel. 23, Created)
DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Putative beta-actin (Fragment).
OS Myxobolus muelleri.
OC Eukaryota; Metazoa; Myxozoa; Myxosporidia; Bivalvulida; Platysoptina;
OC Myxobolidae; Myxobolus.
OX NCBI_TaxID=204749;
RN [1]
RP SEQUENCE FROM N.A.
RA Kelley G.O., Beauchamp K.A., Hedrick R.R.;
RL Submitted (SEP-2002) to the EMBL/GenBank/DBJ databases.
CC -!- FUNCTION: Actins are highly conserved proteins that are involved
CC in various types of cell motility and are ubiquitously expressed
CC in all eukaryotic cells (By similarity).
CC -!- SIMILARITY: Belongs to the actin family.
DR EMBL; AY157010; AAN85095.1; -.
DR HSSP; P02577; 1NM1.
DR GO; GO:0015629; C:actin cytoskeleton; IEA.
DR InterPro; IPR004000; F:structural constituent of cytoskeleton; IEA.
DR Pfam; PF00022; Actin; 1.
DR PRINTS; PR00190; ACTIN.
DR PROSITE; PS01132; ACTINS_ACT_LIKE; 1.
KW Structural protein.
FT NON_TER 1 60
FT NON_TER 60 60
SQ SEQUENCE 60 AA; 6498 MW; 3983D5A1D3D424A7 CRC64;

Query Match 100.0%; Score 74; DB 2; Length 60;
Best Local Similarity 100.0%; Pred. No. 5.5e-06;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 TFNTPMYVAIQAVL 15
Db 26 TFNTPMYVAIQAVL 40

RESULT 7
Q818C4
ID Q818C4 PRELIMINARY; PRT; 60 AA.
AC Q818C4;
DT 01-MAR-2003 (TrEMBLrel. 23, Created)
DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Putative beta-actin (Fragment).
OS Myxobolus muelleri.
OC Eukaryota; Metazoa; Myxozoa; Myxosporidia; Bivalvulida; Platysoptina;
OC Myxobolidae; Myxobolus.
OX NCBI_TaxID=204749;
RN [1]
RP SEQUENCE FROM N.A.
RA Kelley G.O., Beauchamp K.A., Hedrick R.R.;
RL Submitted (SEP-2002) to the EMBL/GenBank/DBJ databases.
CC -!- FUNCTION: Actins are highly conserved proteins that are involved
CC in various types of cell motility and are ubiquitously expressed
CC in all eukaryotic cells (By similarity).
CC -!- SIMILARITY: Belongs to the actin family.
DR EMBL; AY157010; AAN85095.1; -.
DR HSSP; P02577; 1NM1.
DR GO; GO:0015629; C:actin cytoskeleton; IEA.
DR InterPro; IPR004000; F:structural constituent of cytoskeleton; IEA.
DR Pfam; PF00022; Actin; 1.
DR PRINTS; PR00190; ACTIN.
DR PROSITE; PS01132; ACTINS_ACT_LIKE; 1.
KW Structural protein.
FT NON_TER 1 60
FT NON_TER 60 60
SQ SEQUENCE 60 AA; 6498 MW; 3983D5A1D3D424A7 CRC64;

Query Match 100.0%; Score 74; DB 2; Length 60;
Best Local Similarity 100.0%; Pred. No. 5.5e-06;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 TFNTPMYVAIQAVL 15
Db 26 TFNTPMYVAIQAVL 40

RESULT 7
```

```
Q818C5
ID Q818C5 PRELIMINARY; PRT; 60 AA.
AC Q818C5;
DT 01-MAR-2003 (TrEMBLrel. 23, Created)
DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Putative beta-actin (Fragment).
OS Myxobolus spinacuratura.
OC Eukaryota; Metazoa; Myxozoa; Myxosporidia; Bivalvulida; Platysoptina;
OC Myxobolidae; Myxobolus.
OX NCBI_TaxID=182349;
RN [1]
RP SEQUENCE FROM N.A.
RA Kelley G.O., Beauchamp K.A., Hedrick R.R.;
RL Submitted (SEP-2002) to the EMBL/GenBank/DBJ databases.
CC -!- FUNCTION: Actins are highly conserved proteins that are involved
CC in various types of cell motility and are ubiquitously expressed
CC in all eukaryotic cells (By similarity).
CC -!- SIMILARITY: Belongs to the actin family.
DR EMBL; AY157009; AAN85094.1; -.
DR HSSP; P02577; 1NM1.
DR GO; GO:0015629; C:actin cytoskeleton; IEA.
DR InterPro; IPR004000; F:structural constituent of cytoskeleton; IEA.
DR Pfam; PF00022; Actin; 1.
DR PRINTS; PR00190; ACTIN.
DR PROSITE; PS01132; ACTINS_ACT_LIKE; 1.
KW Structural protein.
FT NON_TER 1 60
FT NON_TER 60 60
SQ SEQUENCE 60 AA; 6498 MW; 3983D5A1D3D424A7 CRC64;

Query Match 100.0%; Score 74; DB 2; Length 60;
Best Local Similarity 100.0%; Pred. No. 5.5e-06;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 TFNTPMYVAIQAVL 15
Db 26 TFNTPMYVAIQAVL 40

RESULT 8
Q818C6
ID Q818C6 PRELIMINARY; PRT; 60 AA.
AC Q818C6;
DT 01-MAR-2003 (TrEMBLrel. 23, Created)
DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Putative beta-actin (Fragment).
OS Myxobolus insidiosus.
OC Eukaryota; Metazoa; Myxozoa; Myxosporidia; Bivalvulida; Platysoptina;
OC Myxobolidae; Myxobolus.
OX NCBI_TaxID=59784;
RN [1]
RP SEQUENCE FROM N.A.
RA Kelley G.O., Beauchamp K.A., Hedrick R.R.;
RL Submitted (SEP-2002) to the EMBL/GenBank/DBJ databases.
CC -!- FUNCTION: Actins are highly conserved proteins that are involved
CC in various types of cell motility and are ubiquitously expressed
CC in all eukaryotic cells (By similarity).
CC -!- SIMILARITY: Belongs to the actin family.
DR EMBL; AY157008; AAN85093.1; -.
DR HSSP; P02577; 1NM1.
DR GO; GO:0015629; C:actin cytoskeleton; IEA.
DR InterPro; IPR004000; F:structural constituent of cytoskeleton; IEA.
DR Pfam; PF00022; Actin; 1.
DR PRINTS; PR00190; ACTIN.
DR PROSITE; PS01132; ACTINS_ACT_LIKE; 1.
KW Structural protein.
FT NON_TER 1 60
FT NON_TER 60 60
SQ SEQUENCE 60 AA; 6468 MW; 1483D5A1D3D424B0 CRC64;
```

```
Query Match      100.0%; Score 74; DB 2; Length 60;
Best Local Similarity 100.0%; Pred. No. 5.5e-06;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 TFNTPMYVAIQAVL 15
Db 26 TFNTPMYVAIQAVL 40

RESULT 9
Q6USB7      PRELIMINARY;      PRT;      68 AA.
AC Q6USB7;
DT 05-JUL-2004 (TrEMBLrel. 27, Created)
DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
DT 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
DE Beta-actin (Fragment).
OS Sparus aurata (Gilthead sea bream).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;
OC Acanthomorpha; Acanthopterygii; Percomorpha; Perciformes; Percoidae;
OC Sparidae; Sparus.
OX NCBI_TaxID=8175;
RN [1]
RP SEQUENCE FROM N.A.
RA Krey G.D., Boukouvala E., Hondrogiannis C.I., Theodoridou M.D.,
RA Kalevra V.;
RL Submitted (JUL-2003) to the EMBL/GenBank/DBJ databases.
CC -!- FUNCTION: Actins are highly conserved proteins that are involved
CC in various types of cell motility and are ubiquitously expressed
CC in all eukaryotic cells (By similarity).
CC -!- SIMILARITY: Belongs to the actin family.
DR EMBL: AY362763; AAC56183.1; -.
DR GO: GO:0005629; C:actin cytoskeleton; IEA.
DR GO: GO:0005200; F:structural constituent of cytoskeleton; IEA.
DR InterPro: IPR004000; Actin_like.
DR PRINTS: PR00190; ACTIN.
DR PROSITE: PS01132; ACTINS_ACT_LIKE; 1.
KW Structural protein.
FT NON_TER 1 68
FT NON_TER 68
SQ SEQUENCE 68 AA; 7935 MW; 66E08E30906782D CRC64;

Query Match      100.0%; Score 74; DB 2; Length 68;
Best Local Similarity 100.0%; Pred. No. 6.3e-06;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 TFNTPMYVAIQAVL 15
Db 50 TFNTPMYVAIQAVL 64

RESULT 10
Q8ISR2      PRELIMINARY;      PRT;      95 AA.
AC Q8ISR2;
DT 01-MAR-2003 (TrEMBLrel. 23, Created)
DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Actin (Fragment).
OS Spodoptera frugiperda (Fall armyworm).
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC Neoptera; Endopterygota; Lepidoptera; Glossata; Ditrysia; Noctuoidea;
OC Noctuidae; Amphipyridae; Spodoptera.
OX NCBI_TaxID=7108;
RN [1]
RP SEQUENCE FROM N.A.
RA Nobiron I., O'Reilly D.R., Olszewski J.A.;
RA Submitted (SEP-2002) to the EMBL/GenBank/DBJ databases.
CC -!- FUNCTION: Actins are highly conserved proteins that are involved
CC in various types of cell motility and are ubiquitously expressed
CC in all eukaryotic cells (By similarity).

-!- SIMILARITY: Belongs to the actin family.
EMBL: AF548015; AAN38748.1; -.
HSSP: P02577; INWI
GO: GO:0015629; C:actin cytoskeleton; IEA.
GO: GO:0005200; F:structural constituent of cytoskeleton; IEA.
InterPro: IPR004000; Actin_like.
PFam: PF00022; Actin; 1.
PRINTS: PR00190; ACTIN.
PROSITE: PS01132; ACTINS_ACT_LIKE; 1.
Structural protein.
NON_TER 1 95
NON_TER 95
SEQUENCE 95 AA; 10656 MW; E0C697FD09770932 CRC64;

Query Match      100.0%; Score 74; DB 2; Length 95;
Best Local Similarity 100.0%; Pred. No. 8.8e-06;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 TFNTPMYVAIQAVL 15
Db 46 TFNTPMYVAIQAVL 60

RESULT 11
Q9JLX9      PRELIMINARY;      PRT;      97 AA.
AC Q9JLX9;
DT 01-OCT-2000 (TrEMBLrel. 15, Created)
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE Beta actin (Fragment).
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP SEQUENCE FROM N.A.
RA STRAIN=Sprague-Dawley;
RA Toporsian M., Auer G., Zacour M., Cernacek P., Ward M.E.;
RA Submitted (JAN-1999) to the EMBL/GenBank/DBJ databases.
CC -!- FUNCTION: Actins are highly conserved proteins that are involved
CC in various types of cell motility and are ubiquitously expressed
CC in all eukaryotic cells (By similarity).
CC -!- SIMILARITY: Belongs to the actin family.
EMBL: AF122902; AAF31761.1; -.
HSSP: P02577; INWI.
GO: GO:0015629; C:actin cytoskeleton; IEA.
GO: GO:0005200; F:structural constituent of cytoskeleton; IEA.
InterPro: IPR004000; Actin_like.
PFam: PF00022; Actin; 1.
PRINTS: PR00190; ACTIN.
PROSITE: PS01132; ACTINS_ACT_LIKE; 1.
Structural protein.
NON_TER 1 97
NON_TER 97
SEQUENCE 97 AA; 11014 MW; D5CE6E964444EF9 CRC64;

Query Match      100.0%; Score 74; DB 2; Length 97;
Best Local Similarity 100.0%; Pred. No. 9e-06;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 TFNTPMYVAIQAVL 15
Db 59 TFNTPMYVAIQAVL 73

RESULT 12
Q28916      PRELIMINARY;      PRT;      109 AA.
AC Q28916;
DT 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
```



DE Actin (Fragment).  
 OS Macaca fuscata (Japanese macaque).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Primates; Catarrhini; Cercopithecoidea;  
 OC Cercopithecinae; Macaca.  
 OX NCBI\_TaxID=9542;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=96071113; PubMed=7577171; DOI=10.1016/0960-0760(95)00157-U;  
 RA Yamada-Mouri N., Hirata S., Hayashi M., Kato J.;  
 RT "Analysis of the expression and the first exon of aromatase mRNA in  
 monkey brain.";  
 RL J. Steroid Biochem. Mol. Biol. 55:17-23 (1995).  
 DR EMBL; S79782; AAB35618.2; -.  
 DR HSSP; P10983; 1D4X.  
 DR GO; GO:0015629; C:actin cytoskeleton; IEA.  
 DR GO; GO:0005200; F:structural constituent of cytoskeleton; IEA.  
 DR InterPro; IPR004000; Actin\_like.  
 DR Pfam; PF00022; Actin; 1.  
 FT NON TER 1  
 FT NON TER 109 109  
 SQ SEQUENCE 109 AA; 12333 MW; 6AA64E108CDDA54C CRC64;

Query Match 100.0%; Score 74; DB 2; Length 109;  
 Best Local Similarity 100.0%; Pred. No. 1e-05;  
 Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 TFNTPAMYVAIQAVL 15  
 |||||  
 Db 11 TFNTPAMYVAIQAVL 25

## RESULT 13

Q876P7 ID Q876P7 PRELIMINARY; PRT; 117 AA.  
 AC Q876P7;  
 DT 01-JUN-2002 (TrEMBLrel. 21, Created)  
 DT 01-JUN-2002 (TrEMBLrel. 21, Last sequence update)  
 DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)  
 DE Actin (Fragment).  
 OS Elysia chlorotica (Sea slug).  
 OC Eukaryota; Metazoa; Mollusca; Gastropoda; Orthogastropoda;  
 OC Apogastropoda; Heterobranchia; Euthyneura; Opisthobranchia;  
 OC Sacoglossa; Elysioidea; Elysidae; Elysia.  
 OX NCBI\_TaxID=188477;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA Fukao T., Summer E.J., Rumpho M.E.;  
 RL Submitted (NOV-2001) to the EMBL/GenBank/DBJ databases.  
 CC -!- FUNCTION: Actins are highly conserved proteins that are involved  
 in various types of cell motility and are ubiquitously expressed  
 in all eukaryotic cells (By similarity).  
 CC -!- SIMILARITY: Belongs to the actin family.  
 DR EMBL; AF448493; AAL87639.1; -.  
 DR HSSP; P02577; 1NM1.  
 DR GO; GO:0005884; C:actin filament; IEA.  
 DR GO; GO:0003774; F:motor activity; IEA.  
 DR GO; GO:0005200; F:structural constituent of cytoskeleton; IEA.  
 DR InterPro; IPR004001; Actin.  
 DR Pfam; PF00022; Actin; 1.  
 DR PRINTS; PR00190; ACTIN.  
 DR PROSITE; PS04006; ACTINS\_1; 1.  
 DR PROSITE; PS01132; ACTINS\_ACT\_LIKE; 1.  
 KW Structural protein.  
 FT NON TER 1  
 FT NON TER 117 117  
 SQ SEQUENCE 117 AA; 13270 MW; C5D4DAFDF95F03AC CRC64;

Query Match 100.0%; Score 74; DB 2; Length 117;  
 Best Local Similarity 100.0%; Pred. No. 1.1e-05;  
 Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 TFNTPAMYVAIQAVL 15  
 |||||  
 Db 91 TFNTPAMYVAIQAVL 105

RESULT 14

Q80012 ID Q80012 PRELIMINARY; PRT; 117 AA.  
 AC Q80012;  
 DT 01-JUN-2003 (TrEMBLrel. 24, Created)  
 DT 01-JUN-2003 (TrEMBLrel. 24, Last sequence update)  
 DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)  
 DE Beta-actin 2 (Fragment).  
 GN Name=ACTB2;  
 OS Hyla japonica (Japanese tree frog).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Amphibia; Batrachia; Anura; Neobatrachia; Hyloidea; Hylinae;  
 OC Hyla.  
 OX NCBI\_TaxID=109175;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=Liver;  
 RA Kohno S., Kamishima Y., Iguchi T.;  
 RL Submitted (SEP-2002) to the EMBL/GenBank/DBJ databases.  
 CC -!- FUNCTION: Actins are highly conserved proteins that are involved  
 in various types of cell motility and are ubiquitously expressed  
 in all eukaryotic cells (By similarity).  
 CC -!- SIMILARITY: Belongs to the actin family.  
 DR EMBL; AB092520; BAC66208.1; -.  
 DR HSSP; P60712; 1HLU.  
 DR GO; GO:0015629; C:actin cytoskeleton; IEA.  
 DR GO; GO:0005200; F:structural constituent of cytoskeleton; IEA.  
 DR InterPro; IPR004000; Actin\_like.  
 DR Pfam; PF00022; Actin; 1.  
 DR PRINTS; PR00190; ACTIN.  
 DR PROSITE; PS01132; ACTINS\_ACT\_LIKE; 1.  
 KW Structural protein.  
 FT NON TER 1  
 FT NON TER 117 117  
 SQ SEQUENCE 117 AA; 13291 MW; 64C6BAC4F84184CF CRC64;

Query Match 100.0%; Score 74; DB 2; Length 117;  
 Best Local Similarity 100.0%; Pred. No. 1.1e-05;  
 Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 TFNTPAMYVAIQAVL 15  
 |||||  
 Db 47 TFNTPAMYVAIQAVL 61

RESULT 15

Q80013 ID Q80013 PRELIMINARY; PRT; 117 AA.  
 AC Q80013;  
 DT 01-JUN-2003 (TrEMBLrel. 24, Created)  
 DT 01-JUN-2003 (TrEMBLrel. 24, Last sequence update)  
 DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)  
 DE Beta-actin 1 (Fragment).  
 GN Name=ACTB1;  
 OS Hyla japonica (Japanese tree frog).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Amphibia; Batrachia; Anura; Neobatrachia; Hyloidea; Hylinae;  
 OC Hyla.  
 OX NCBI\_TaxID=109175;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=Liver;  
 RA Kohno S., Kamishima Y., Iguchi T.;  
 RL Submitted (SEP-2002) to the EMBL/GenBank/DBJ databases.  
 CC -!- FUNCTION: Actins are highly conserved proteins that are involved  
 in various types of cell motility and are ubiquitously expressed  
 in all eukaryotic cells (By similarity).  
 CC -!- SIMILARITY: Belongs to the actin family.

Query Match 100.0%; Score 74; DB 2; Length 117;  
 Best Local Similarity 100.0%; Pred. No. 1.1e-05;  
 Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 TFNTPAMYVAIQAVL 15  
 |||||  
 Db 47 TFNTPAMYVAIQAVL 61

RESULT 15

Q80013 ID Q80013 PRELIMINARY; PRT; 117 AA.  
 AC Q80013;  
 DT 01-JUN-2003 (TrEMBLrel. 24, Created)  
 DT 01-JUN-2003 (TrEMBLrel. 24, Last sequence update)  
 DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)  
 DE Beta-actin 1 (Fragment).  
 GN Name=ACTB1;  
 OS Hyla japonica (Japanese tree frog).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Amphibia; Batrachia; Anura; Neobatrachia; Hyloidea; Hylinae;  
 OC Hyla.  
 OX NCBI\_TaxID=109175;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=Liver;  
 RA Kohno S., Kamishima Y., Iguchi T.;  
 RL Submitted (SEP-2002) to the EMBL/GenBank/DBJ databases.  
 CC -!- FUNCTION: Actins are highly conserved proteins that are involved  
 in various types of cell motility and are ubiquitously expressed  
 in all eukaryotic cells (By similarity).  
 CC -!- SIMILARITY: Belongs to the actin family.

DR EMBL; AB092519; BAC66207.1; -.  
 DR HSSP; P60712; 1HLU.  
 DR GO; GO:0015629; C:actin cytoskeleton; IEA.  
 DR GO; GO:0005200; F:structural constituent of cytoskeleton; IEA.  
 DR InterPro; IPR004000; Actin\_like.  
 DR Pfam; PF00022; Actin; 1.  
 DR PRINTS; PR00190; ACTIN.  
 DR PROSITE; PS01132; ACTINS\_ACT\_LIKE; 1.  
 KW Structural protein.  
 FT NON\_TER 1  
 FT NON\_TER 117  
 SQ SEQUENCE 117 AA; 13291 MW; 64C6BAC4F84184CF CRC64;

Query Match 100.0%; Score 74; DB 2; Length 117;  
 Best Local Similarity 100.0%; Pred. No. 1.1e-05;  
 Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 TFNTPAMYVAIQAVL 15  
 |||||  
 Db 47 TFNTPAMYVAIQAVL 61

Search completed: April 8, 2005, 12:03:09  
 Job time : 58.8667 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: April 8, 2005, 09:58:56 ; Search time 62.4 Seconds  
(without alignments)  
92.971 Million cell updates/sec

Title: US-09-423-351C-7

Perfect score: 75

Sequence: 1 LFHAILRLDLAGRDL 15

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 2105692 seqs, 386760381 residues

Total number of hits satisfying chosen parameters: 2105692

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Listing first 45 summaries

Database : A\_Geneseq\_16Dec04:\*

- 1: geneseqp1980s:\*
- 2: geneseqp1990s:\*
- 3: geneseqp2000s:\*
- 4: geneseqp2001s:\*
- 5: geneseqp2002s:\*
- 6: geneseqp2003as:\*
- 7: geneseqp2003bs:\*
- 8: geneseqp2004s:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	75	100.0	15	2	Aaw92533 Beta-acti
2	75	100.0	97	4	Aam13687 Peptide #
3	75	100.0	97	4	Abb32618 Peptide #
4	75	100.0	97	4	Aam26087 Peptide #
5	75	100.0	97	4	Abb27467 Human pep
6	75	100.0	97	4	Abb18116 Protein #
7	75	100.0	97	4	Aam65825 Human bon
8	75	100.0	97	4	Aam53448 Human bra
9	75	100.0	97	4	Abg47471 Human liv
10	75	100.0	97	4	Aam01437 Peptide #
11	75	100.0	97	5	Abg35459 Human pep
12	75	100.0	118	3	Aag26468 Arabidops
13	75	100.0	135	3	Aag26467 Arabidops
14	75	100.0	148	5	Adk36534 Novel hum
15	75	100.0	150	4	Abg15101 Novel hum
16	75	100.0	153	4	Abg15497 Novel hum
17	75	100.0	168	5	Abp42912 Human ova
18	75	100.0	196	6	Abu70816 Human adi
19	75	100.0	211	7	Abm73899 DNA clone
20	75	100.0	229	3	Aag38700 Arabidops
21	75	100.0	246	3	Aag38699 Arabidops
22	75	100.0	257	3	Aag37992 Arabidops
23	75	100.0	257	3	Aag20916 Arabidops
24	75	100.0	274	6	Abu70549 Human adi
25	75	100.0	276	3	Aag50947 Arabidops

## ALIGNMENTS

## RESULT 1

Aaw92533  
ID Aaw92533 standard; peptide; 15 AA.

AC Aaw92533;

XX  
DT 26-APR-1999 (first entry)

DE Beta-actin reference peptide substrate #7.

KW Peptide substrate; CCT; eukaryotic type II chaperonin complex; cyclin; binding agent; substrate-binding site; SBS; substrate folding; actin; tubulin; treatment; cancer; anticancer drug; viral infection; screening; reduced toxicity.

OS Synthetic.

PN WO9853322-A1.

PD 26-NOV-1998.

PF 22-MAY-1998; 98WO-GB001485.

PR 23-MAY-1997; 97GB-00010762.

PA (CANC-) INST CANCER RES ROYAL CANCER HOSPITAL.

PI Willison K, Hynes G, Liou AK;

XX WPI; 1999-070162/06.

PT Identifying specific binding agents for substrate binding site in CCT chaperonin complex - also new peptide binding agents and their mimetics, and peptides containing a specific CCT binding site, used for treating cancer.

PS Disclosure; Fig 10; 97pp; English.

CC This invention describes a method which uses the CCT (eukaryotic type II chaperonin) complex or part of it, for identifying a binding agent that can occupy a substrate-binding site (SBS) on the CCT complex. By binding to the CCT complex, the binding agents block an SBS so that biological activity of the CCT complex is affected, particularly its ability to fold substrates such as actin, tubulin and cyclin. The binding agents are useful for treatment of cancer, particularly when used in combination with an anticancer drug, or viral infections. Nucleic acid fragments are used to screen for agents, e.g. binding agents that modulate interaction between the CCT complex and a protein that is to be folded. The binding

Aag05115 Arabidops  
Aag20915 Arabidops  
Aag37991 Arabidops  
Adi63068 Human apo  
Aag20914 Arabidops  
Aag37990 Arabidops  
Aag50946 Arabidops  
Aag05114 Arabidops  
Aag37210 Arabidops  
Aag37209 Arabidops  
Aag50945 Arabidops  
Aag05113 Arabidops  
Aag06448 Arabidops  
Aag30774 Arabidops  
Aag50873 Arabidops  
Aag07390 Arabidops  
Aag41511 Arabidops  
Aag37989 Arabidops  
Aag37317 Arabidops  
Aag44697 Zea mays

26 75 100.0 277 3 AAG05115  
27 75 100.0 294 3 AAG20915  
28 75 100.0 294 3 AAG37991  
29 75 100.0 294 7 ADI63068  
30 75 100.0 295 3 AAG20914  
31 75 100.0 295 3 AAG37990  
32 75 100.0 304 3 AAG50946  
33 75 100.0 305 3 AAG05114  
34 75 100.0 308 3 AAG37210  
35 75 100.0 311 3 AAG37209  
36 75 100.0 321 3 AAG50945  
37 75 100.0 322 3 AAG05113  
38 75 100.0 332 3 AAG06448  
39 75 100.0 332 3 AAG30774  
40 75 100.0 332 3 AAG50873  
41 75 100.0 332 3 AAG07390  
42 75 100.0 332 3 AAG41511  
43 75 100.0 332 3 AAG37989  
44 75 100.0 332 3 AAG37317  
45 75 100.0 332 3 AAG44697

CC agents may target cells that are actively synthesising tubulin etc.  
CC (unlike known microtubule-stabilising agents that affect all cells), so  
CC should have reduced toxicity for normal cells. AAW92527-W92541 are  
CC peptide substrates used in the method of the invention  
XX  
XX  
SQ Sequence 15 AA;

Query Match 100.0%; Score 75; DB 2; Length 15;  
Best Local Similarity 100.0%; Pred. No. 2.2e-07;  
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 LPHAILRLDLGRDL 15  
Db 1 LPHAILRLDLGRDL 15  
|||||

RESULT 2  
AA113687  
ID AA113687 standard; protein; 97 AA.

XX  
AC AA113687;

XX  
DT 12-OCT-2001 (first entry)

XX Peptide #121 encoded by probe for measuring cervical gene expression.  
DE Probe; human; microarray; gene expression; cervical epithelial cell;  
XX cervical cancer.

XX Homo sapiens.

XX WO200157278-A2.

XX  
PD 09-AUG-2001.

XX  
PF 30-JAN-2001; 2001WO-US000670.

XX  
PR 04-FEB-2000; 2000US-0180312P.

PR 26-MAY-2000; 2000US-0207456P.

PR 30-JUN-2000; 2000US-00608408.

PR 03-AUG-2000; 2000US-00632366.

PR 21-SEP-2000; 2000US-0234687P.

PR 27-SEP-2000; 2000US-0236359P.

PR 04-OCT-2000; 2000GB-00024263.

XX (MOLE-) MOLECULAR DYNAMICS INC.

XX Penn SG, Hanzel DK, Chen W, Rank DR;

XX WPI; 2001-488901/53.

XX Claim 27; SEQ ID NO 18513; 487pp; English.

XX The present invention relates to human single exon nucleic acid probes  
CC (SENPs: see AA110068-AA128459). The present sequence is a peptide encoded  
CC by one such probe. The SENPs are derived from human HeLa cells. The SENPs  
CC can be used to produce a single exon microarray, which can be used for  
CC measuring human gene expression in a sample derived from human cervical  
CC epithelial cells. By measuring gene expression, the probes are therefore  
CC useful in grading and/or staging of diseases of the cervix, notably  
CC cervical cancer. Note: The sequence data for this patent did not form  
CC part of the printed specification, but was obtained in electronic format  
CC directly from WIPO at ftp.wipo.int/pub/published\_pct\_sequences

XX Sequence 97 AA;

Query Match 100.0%; Score 75; DB 4; Length 97;  
Best Local Similarity 100.0%; Pred. No. 2.1e-06;  
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 LPHAILRLDLGRDL 15  
Db 50 LPHAILRLDLGRDL 64  
|||||

RESULT 3  
ABB32618

ID ABB32618 standard; peptide; 97 AA.

XX  
AC ABB32618;

XX  
DT 04-FEB-2002 (first entry)

XX Peptide #124 encoded by human foetal liver single exon probe.

XX Human; foetal liver; gene expression; single exon nucleic acid probe.

XX Homo sapiens.

XX WO200157277-A2.

XX  
PD 09-AUG-2001.

XX  
PF 30-JAN-2001; 2001WO-US000669.

XX  
PR 04-FEB-2000; 2000US-0180312P.

PR 26-MAY-2000; 2000US-0207456P.

PR 30-JUN-2000; 2000US-00608408.

PR 03-AUG-2000; 2000US-00632366.

PR 21-SEP-2000; 2000US-0234687P.

PR 27-SEP-2000; 2000US-0236359P.

PR 04-OCT-2000; 2000GB-00024263.

XX (MOLE-) MOLECULAR DYNAMICS INC.

XX Penn SG, Hanzel DK, Chen W, Rank DR;

XX WPI; 2001-483447/52.

XX Human genome-derived single exon nucleic acid probes useful for analyzing

XX gene expression in human fetal liver.

XX Claim 27; SEQ ID NO 25253; 639pp + Sequence Listing; English.

XX The invention relates to a single exon nucleic acid probe for measuring

XX human gene expression in a sample derived from human foetal liver. The

XX single exon nucleic acid probes may be used for predicting, measuring and

XX displaying gene expression in samples derived from human fetal liver. The

XX present sequence is a peptide encoded by a single exon nucleic acid probe

XX of the invention. Note: The sequence data for this patent did not form

XX part of the printed specification, but was obtained in electronic format

XX directly from WIPO at ftp.wipo.int/pub/published\_pct\_sequences

XX Sequence 97 AA;

Query Match 100.0%; Score 75; DB 4; Length 97;  
Best Local Similarity 100.0%; Pred. No. 2.1e-06;  
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 LPHAILRLDLGRDL 15  
Db 50 LPHAILRLDLGRDL 64  
|||||

RESULT 4  
AA126087

ID AA126087 standard; protein; 97 AA.

XX  
AC AA126087;

XX  
DT 17-OCT-2001 (first entry)

XX Peptide #124 encoded by probe for measuring placental gene expression.

XX Probe; microarray; human; placenta; antenatal diagnosis;  
KW genetic disorder.

XX Homo sapiens.

XX WO200157272-A2.

XX 09-AUG-2001.

XX 30-JAN-2001; 2001WO-US000663.

XX 04-FEB-2000; 2000US-0180312P.

XX 26-MAY-2000; 2000US-0207456P.

XX 30-JUN-2000; 2000US-00608408.

XX 03-AUG-2000; 2000US-00632366.

XX 21-SEP-2000; 2000US-0234687P.

XX 27-SEP-2000; 2000US-0236359P.

XX 04-OCT-2000; 2000GB-00024263.

XX (MOLE-) MOLECULAR DYNAMICS INC.

XX Penn SG, Hanzel DK, Chen W, Rank DR;

XX WPI; 2001-488897/53.

XX Human genome-derived single exon nucleic acid probes useful for analyzing

XX gene expression in human placenta.

XX Claim 27; SEQ ID NO 26356; 654pp; English.

XX The present invention relates to single exon nucleic acid probes (SENPs;

XX see AAT31315-AA157546). The present sequence is a peptide encoded by one

XX such probe. The probes are useful for producing a microarray for

XX predicting, measuring and displaying gene expression in samples derived

XX from human placenta. The probes are useful for antenatal diagnosis of

XX human genetic disorders

XX Sequence 97 AA;

XX Query Match 100.0%; Score 75; DB 4; Length 97;

XX Best Local Similarity 100.0%; Pred. No. 2.1e-06;

XX Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

XX 1 LPHAILRLDLAGRDL 15

XX |||||

XX 50 LPHAILRLDLAGRDL 64

XX RESULT 5

XX ABB27467

XX ID ABB27467 standard; peptide; 97 AA.

XX AC ABB27467;

XX 01-FEB-2002 (first entry)

XX Human peptide #118 encoded by breast cell single exon nucleic acid probe.

XX Human; microarray; single exon probe; gene expression; breast; disease;

XX cancer.

XX Homo sapiens.

XX WO200157271-A2.

XX 09-AUG-2001.

XX 30-JAN-2001; 2001WO-US000662.

XX 04-FEB-2000; 2000US-0180312P.

XX 26-MAY-2000; 2000US-0207456P.

XX 30-JUN-2000; 2000US-00608408.

XX 03-AUG-2000; 2000US-00632366.

XX 21-SEP-2000; 2000US-0234687P.

XX 27-SEP-2000; 2000US-0236359P.

XX 04-OCT-2000; 2000GB-00024263.

XX (MOLE-) MOLECULAR DYNAMICS INC.

XX Penn SG, Hanzel DK, Chen W, Rank DR;

XX WPI; 2001-488897/53.

XX Human genome-derived single exon nucleic acid probes useful for analyzing

XX gene expression in human placenta.

XX Claim 27; SEQ ID NO 26356; 654pp; English.

XX The present invention relates to single exon nucleic acid probes (SENPs;

XX see AAT31315-AA157546). The present sequence is a peptide encoded by one

XX such probe. The probes are useful for producing a microarray for

XX predicting, measuring and displaying gene expression in samples derived

XX from human placenta. The probes are useful for antenatal diagnosis of

XX human genetic disorders

XX Sequence 97 AA;

XX Query Match 100.0%; Score 75; DB 4; Length 97;

XX Best Local Similarity 100.0%; Pred. No. 2.1e-06;

XX Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

XX 1 LPHAILRLDLAGRDL 15

XX |||||

XX 50 LPHAILRLDLAGRDL 64

XX RESULT 5

XX ABB27467

XX ID ABB27467 standard; peptide; 97 AA.

XX AC ABB27467;

XX 01-FEB-2002 (first entry)

XX Human peptide #118 encoded by breast cell single exon nucleic acid probe.

XX Human; microarray; single exon probe; gene expression; breast; disease;

XX cancer.

XX Homo sapiens.

XX WO200157272-A2.

XX 09-AUG-2001.

XX 30-JAN-2001; 2001WO-US000663.

XX 04-FEB-2000; 2000US-0180312P.

XX 26-MAY-2000; 2000US-0207456P.

XX 30-JUN-2000; 2000US-00608408.

XX 03-AUG-2000; 2000US-00632366.

XX 21-SEP-2000; 2000US-0234687P.

XX 27-SEP-2000; 2000US-0236359P.

XX 04-OCT-2000; 2000GB-00024263.

XX (MOLE-) MOLECULAR DYNAMICS INC.

XX Penn SG, Hanzel DK, Chen W, Rank DR;

XX WPI; 2001-496933/54.

XX New spatially-addressable set of single exon nucleic acid probes, useful

XX for measuring gene expression in sample derived from human breast,

XX comprises number of single exon nucleic acid probes.

XX Claim 27; SEQ ID NO 10435; 327pp + Sequence Listing; English.

XX The invention relates to a spatially-addressable set of single exon

XX nucleic acid probes for measuring gene expression in a sample derived

XX from human breast and BT 474 cells. The method involves contacting the

XX probes with a collection of detectably labelled nucleic acids derived

XX from mRNA of human breast, and then measuring the label bound to each

XX probe of the microarray. The probes are useful for verifying the

XX expression of regions of genomic DNA predicted to encode proteins. They

XX are useful for gene discovery, and for determining predisposition and/or

XX assessing breast disease. Gene expression analysis is useful for

XX this invention presents a far greater diversity of probes for measuring

XX gene expression, with far less bias than expressed sequence tag

XX microarrays. The method is suitable for rapid production of functional

XX information from genomic sequence. The present sequence is a peptide

XX encoded by a single exon nucleic acid probe of the invention. Note: The

XX sequence data for this patent did not form part of the printed

XX specification, but was obtained in electronic format directly from WIPO

XX at ftp.wipo.int/pub/published\_pct\_sequences

XX Sequence 97 AA;

XX Query Match 100.0%; Score 75; DB 4; Length 97;

XX Best Local Similarity 100.0%; Pred. No. 2.1e-06;

XX Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

XX 1 LPHAILRLDLAGRDL 15

XX |||||

XX 50 LPHAILRLDLAGRDL 64

XX RESULT 6

XX ABB18116

XX ID ABB18116 standard; protein; 97 AA.

XX AC ABB18116;

XX 23-JAN-2002 (first entry)

XX Protein #115 encoded by probe for measuring heart cell gene expression.

XX Human; gene expression; heart; microarray; vascular system;

XX cardiovascular disease; hypertension; cardiac arrhythmia;

XX congenital heart disease.

XX Homo sapiens.

XX WO200157274-A2.

XX 09-AUG-2001.

XX 30-JAN-2001; 2001WO-US000666.

XX 04-FEB-2000; 2000US-0180312P.

XX 26-MAY-2000; 2000US-0207456P.

XX 30-JUN-2000; 2000US-00608408.

XX 03-AUG-2000; 2000US-00632366.

```
PR 21-SEP-2000; 2000US-0234687P.
PR 27-SEP-2000; 2000US-0236359P.
PR 04-OCT-2000; 2000GB-00024263.
XX
PA (MOLE-) MOLECULAR DYNAMICS INC.
XX
PI Penn SG, Hanzel DK, Chen W, Rank DR;
XX
DR WPI; 2001-488999/53.
XX
PT Single exon nucleic acid probes for analyzing gene expression in human
PT hearts.
XX
PS Claim 15; SEQ ID NO 19886; 530pp; English.
XX
CC The present invention relates to single exon nucleic acid probes for
CC measuring human gene expression in a sample derived from human heart (see
CC ABA21535-ABA41305). The present sequence is a protein encoded by one such
CC probe. The probes may be used for predicting, measuring and displaying
CC gene expression in samples derived from the human heart via microarrays.
CC By measuring gene expression, the probes are useful for predicting,
CC diagnosing, grading, staging, monitoring and prognosing diseases of the
CC human heart and vascular system e.g. cardiovascular disease,
CC hypertension, cardiac arrhythmias and congenital heart disease. Note: The
CC sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at ftp.wipo.int/pub/published_pct_sequences
XX
XX Sequence 97 AA;
XX
Query Match 100.0%; Score 75; DB 4; Length 97;
Best Local Similarity 100.0%; Pred. No. 2.1e-06;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 LPFAILRLDLGRDL 15
Db 50 LPFAILRLDLGRDL 64
RESULT 7
AAM65825
ID AAM65825 standard; protein; 97 AA.
XX
AC AAM65825;
XX
DT 06-NOV-2001 (first entry)
XX
DE Human bone marrow expressed probe encoded protein SEQ ID NO: 26131.
XX
KW Human; bone marrow expressed exon; gene expression analysis; probe;
KW microarray; cancer; leukaemia; lymphoma; myeloma.
XX
OS Homo sapiens.
XX
PN WO200157276-A2.
XX
PD 09-AUG-2001.
XX
PF 30-JAN-2001; 2001WO-US0000668.
XX
PR 04-FEB-2000; 2000US-0180312P.
PR 26-MAY-2000; 2000US-0207456P.
PR 30-JUN-2000; 2000US-00608408.
PR 03-AUG-2000; 2000US-00632366.
PR 21-SEP-2000; 2000US-0234687P.
PR 27-SEP-2000; 2000US-0236359P.
PR 04-OCT-2000; 2000GB-00024263.
XX
PA (MOLE-) MOLECULAR DYNAMICS INC.
XX
PI Penn SG, Hanzel DK, Chen W, Rank DR;
XX
DR WPI; 2001-488900/53.
XX
XX Sequence 97 AA;
XX
Query Match 100.0%; Score 75; DB 4; Length 97;
Best Local Similarity 100.0%; Pred. No. 2.1e-06;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 LPFAILRLDLGRDL 15
Db 50 LPFAILRLDLGRDL 64
RESULT 8
AAM53448
ID AAM53448 standard; protein; 97 AA.
XX
AC AAM53448;
XX
DT 05-NOV-2001 (first entry)
XX
DE Human brain expressed single exon probe encoded protein SEQ ID NO: 25553.
XX
KW Human; brain expressed exon; gene expression analysis; probe; microarray;
KW Alzheimer's disease; multiple sclerosis; schizophrenia; epilepsy; cancer.
XX
OS Homo sapiens.
XX
PN WO200157275-A2.
XX
PD 09-AUG-2001.
XX
PF 30-JAN-2001; 2001WO-US0000667.
XX
PR 04-FEB-2000; 2000US-0180312P.
PR 26-MAY-2000; 2000US-0207456P.
PR 30-JUN-2000; 2000US-00608408.
PR 03-AUG-2000; 2000US-00632366.
PR 21-SEP-2000; 2000US-0234687P.
PR 27-SEP-2000; 2000US-0236359P.
PR 04-OCT-2000; 2000GB-00024263.
XX
PA (MOLE-) MOLECULAR DYNAMICS INC.
XX
PI Penn SG, Hanzel DK, Chen W, Rank DR;
XX
DR WPI; 2001-483446/52.
XX
XX Single exon nucleic acid probes for analyzing gene expression in human
XX brains.
XX
PS Example 4; SEQ ID NO 25553; 650pp + Sequence Listing; English.
XX
CC The present invention provides a number of single exon nucleic acid
CC probes which are derived from genomic sequences expressed in the human
CC brain. They can be used to measure gene expression in brain cell samples,
CC which may enable the diagnosis and improved treatment of nervous system
CC diseases such as Alzheimer's disease, multiple sclerosis, schizophrenia,
CC epilepsy and cancers. The present sequence is a protein encoded by one of
CC the probes of the invention
XX
XX Sequence 97 AA;
```

Query Match 100.0%; Score 75; DB 4; Length 97;  
Best Local Similarity 100.0%; Pred. No. 2.1e-06;  
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 LPHAILRLDLGRDL 15  
|||||  
Db 50 LPHAILRLDLGRDL 64

RESULT 9  
ID ABG47471 standard; peptide; 97 AA.

XX AC ABG47471;  
XX XX  
DT 25-FEB-2003 (first entry)  
XX XX  
DE Human liver peptide, SEQ ID No 26119.  
XX XX  
KW Human; liver; cirrhosis; hyperlipoproteinaemia; hyperlipidaemia;  
KW hypercholesterolaemia; coronary heart disease.  
XX XX

OS Homo sapiens.

XX XX  
PN WO200157273-A2.

XX XX  
PD 09-AUG-2001.

XX XX  
PF 30-JAN-2001; 2001WO-US000664.

XX XX  
PR 04-FEB-2000; 2000US-0180312P.

XX XX  
PR 26-MAY-2000; 2000US-0207456P.

XX XX  
PR 30-JUN-2000; 2000US-00608408.

XX XX  
PR 03-AUG-2000; 2000US-00632386.

XX XX  
PR 21-SEP-2000; 2000US-0234687P.

XX XX  
PR 27-SEP-2000; 2000US-0236359P.

XX XX  
PR 04-OCT-2000; 2000GB-00024263.

XX (MOLE-) MOLECULAR DYNAMICS INC.

XX PA Penn SG, Hanzel DK, Chen W, Rank DR;

XX PI WPI; 2001-488898/53.

XX XX  
PT Human genome-derived single exon nucleic acid probes useful for analyzing  
PT gene expression in human adult liver.

XX PS Claim 27; SEQ ID NO 26119; 658pp; English.

XX CC The invention relates to a single exon nucleic acid probe (SENP) (I) for  
CC measuring human gene expression in a sample derived from human adult  
CC liver, comprising one of 13109 defined nucleotide sequences given in the  
CC specification (or complements/ fragments). The probe hybridizes at high  
CC stringency to a nucleic acid molecule expressed in the human adult liver.  
CC (I) may be used for predicting, measuring and displaying gene expression  
CC in samples derived from human adult liver. The genes identified may be  
CC involved in genetic liver diseases such as cirrhosis,  
CC hyperlipoproteinaemia, hyperlipidaemia and hypercholesterolaemia which is  
CC associated with coronary heart disease. ABG47348-ABG5930 represent human  
CC liver single exon encoded peptides of the invention. Note: The sequence  
CC information for this patent does not appear in the printed specification  
CC but was obtained in electronic format directly from WIPO at  
CC ftp.wipo.int/pub/published\_pct\_sequences

XX SQ Sequence 97 AA;

Query Match 100.0%; Score 75; DB 4; Length 97;  
Best Local Similarity 100.0%; Pred. No. 2.1e-06;  
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 LPHAILRLDLGRDL 15  
|||||

Db 50 LPHAILRLDLGRDL 64

RESULT 10  
AAM01437

ID AAM01437 standard; protein; 97 AA.

XX AC AAM01437;

XX XX  
DT 09-OCT-2001 (first entry)

XX XX  
DE Peptide #119 encoded by probe for measuring human breast gene expression.  
XX XX  
KW Probe; human; breast disease; breast cancer; development disorder;  
KW inflammatory disease; proliferative breast disease; non-carcinoma tumour.

XX OS Homo sapiens.

XX XX  
PN WO200157270-A2.

XX XX  
PD 09-AUG-2001.

XX XX  
PF 29-JAN-2001; 2001WO-US000661.

XX XX  
PR 04-FEB-2000; 2000US-0180312P.

XX XX  
PR 26-MAY-2000; 2000US-0207456P.

XX XX  
PR 30-JUN-2000; 2000US-00608408.

XX XX  
PR 03-AUG-2000; 2000US-00632366.

XX XX  
PR 21-SEP-2000; 2000US-0234687P.

XX XX  
PR 27-SEP-2000; 2000US-0236359P.

XX XX  
PR 04-OCT-2000; 2000GB-00024263.

XX (MOLE-) MOLECULAR DYNAMICS INC.

XX XX  
PI Penn SG, Hanzel DK, Chen W, Rank DR;

XX XX  
DR WPI; 2001-476286/51.

XX XX  
PT Novel single exon nucleic acid probe used to measuring gene expression in  
PT a human breast.

XX PS Claim 27; SEQ ID NO 10177; 322pp; English.

XX CC The present invention relates to novel single exon nucleic acid probes  
CC (see AAI00010-AA110067). The present sequence is a peptide encoded by one  
CC such probe. The probes are useful for measuring human gene expression in  
CC a human breast sample, where the probe hybridizes at high stringency to a  
CC nucleic acid expressed in the human breast. The probes are useful for  
CC predicting, diagnosing, grading, staging, monitoring and prognosing  
CC diseases of the human breast, particularly those diseases with polygenic  
CC aetiology. The diseases include: breast cancer, disorders of development,  
CC inflammatory diseases of the breast, fibrocystic changes, proliferative  
CC breast disease and non-carcinoma tumours. Note: The sequence data for  
CC this patent did not form part of the printed specification, but was  
CC obtained in electronic format directly from WIPO at  
CC ftp.wipo.int/pub/published\_pct\_sequences

XX SQ Sequence 97 AA;

Query Match 100.0%; Score 75; DB 4; Length 97;  
Best Local Similarity 100.0%; Pred. No. 2.1e-06;  
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 LPHAILRLDLGRDL 15  
|||||  
Db 50 LPHAILRLDLGRDL 64

RESULT 11  
ABG35459

ID ABG35459 standard; peptide; 97 AA.

XX AC ABG35459;

XX 19-AUG-2002 (first entry)

XX Human peptide encoded by genome-derived single exon probe SEQ ID 25124.

DE Human; single exon probe; asthma; lung cancer; COPD; ILD;

XX chronic obstructive pulmonary disease; interstitial lung disease;

XX familial idiopathic pulmonary fibrosis; neurofibromatosis;

XX tuberous sclerosis; Gaucher's disease; Niemann-Pick disease;

XX Hermansky-Pudlak syndrome; sarcoidosis; pulmonary haemosiderosis;

XX pulmonary histiocytosis; lymphangioleiomyomatosis; Karagener syndrome;

XX pulmonary alveolar proteinosis; fibrocystic pulmonary dysplasia;

XX primary ciliary dyskinesia; pulmonary hypertension;

XX hyaline membrane disease.

XX Homo sapiens.

XX WO200186003-A2.

XX 15-NOV-2001.

XX 30-JAN-2001; 2001WO-US0000665.

XX 04-FEB-2000; 2000US-0180312P.

XX 26-MAY-2000; 2000US-0207456P.

XX 30-JUN-2000; 2000US-00608408.

XX 03-AUG-2000; 2000US-00632366.

XX 21-SEP-2000; 2000US-0234687P.

XX 27-SEP-2000; 2000US-0236359P.

XX 04-OCT-2000; 2000GB-00024263.

XX (MOLE-) MOLECULAR DYNAMICS INC.

XX Penn SG, Hanzel DK, Chen W, Rank DR;

XX WPI; 2002-114183/15.

XX Spatially-addressable set of single exon nucleic acid probes, used to

XX measure gene expression in human lung samples.

XX Claim 27; SEQ ID NO 25124; 634pp; English.

XX The invention relates to a spatially-addressable set of single exon

XX nucleic acid probes for measuring gene expression in a sample derived

XX from human lung comprising single exon nucleic acid probes having one of

XX 12614 nucleic acid sequences mentioned in the specification, or their

XX complements or the 12387 open reading frames derived from the 12614

XX probes. Also included are a microarray comprising the novel set of probes

XX; the novel set of probes which hybridize at high stringency to a nucleic

XX acid expressed in the human lung; measuring gene expression in a sample

XX derived from human lung, comprising (a) contacting the array with a

XX collection of detectably labeled nucleic acids derived from human lung

XX mRNA, and (b) measuring the label detectably bound to each probe of the

XX array; identifying exons in a eukaryotic genome, comprising (a)

XX algorithmically predicting at least one exon from genomic sequences of

XX the eukaryote; and (b) detecting specific hybridization of detectably

XX labeled nucleic acids from eukaryote lung mRNA, to a single exon probe,

XX having a fragment identical to the predicted exon, the probe is included

XX in the above mentioned microarray; assigning exons to a single gene,

XX comprising (a) identifying exons from genomic sequence by the method

XX above and (b) measuring the expression of each of the exons in several

XX tissues and/or cell types using hybridization to a single exon

XX microarrays having a probe with the exon, where a common pattern of

XX expression of the exons in the tissues and/or cell types indicates that

XX the exons should be assigned to a single gene; a peptide comprising one

XX of 12011 sequences, mentioned in the specification, or encoded by the

XX probes/open reading frames (ORF). The probes are used for gene expression

XX analysis, and for identifying exons in a gene, particularly using human

XX lung derived mRNA and for the study of lung diseases such as asthma, lung

XX cancer, chronic obstructive pulmonary disease (COPD), interstitial lung

XX disease (ILD), familial idiopathic pulmonary fibrosis, neurofibromatosis,

XX tuberous sclerosis, Gaucher's disease, Niemann-Pick disease, Hermansky-

XX Pudlak syndrome, sarcoidosis, pulmonary haemosiderosis, pulmonary

CC histiocytosis, lymphangioleiomyomatosis, pulmonary alveolar proteinosis,

CC Karagener syndrome, fibrocystic pulmonary dysplasia, primary ciliary

CC dyskinesia, pulmonary hypertension and hyaline membrane disease. The

CC present sequence is a peptide/protein encoded by a single exon probe of

CC the invention. Note: The sequence data for this patent did not form part

CC of the printed specification, but was obtained in electronic format

CC directly from WIPO at ftp.wipo.int/pub/published\_pct\_sequences

XX Sequence 97 AA;

XX Query Match 100.0%; Score 75; DB 5; Length 97;

XX Best Local Similarity 100.0%; Pred. No. 2.1e-06;

XX Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 LPHAILRLDLAARDL 15

DB 50 LPHAILRLDLAARDL 64

RESULT 12

AG26468

ID AAG26468 standard; protein; 118 AA.

XX AC AAG26468;

XX DT 17-OCT-2000 (first entry)

XX DE Arabidopsis thaliana protein fragment SEQ ID NO: 30934.

XX KW Protein identification; signal transduction pathway; metabolic pathway;

XX KW hybridisation assay; genetic mapping; gene expression control; promoter;

XX KW termination sequence.

XX OS Arabidopsis thaliana.

XX EP1033405-A2.

PD 06-SEP-2000.

XX 25-FEB-2000; 2000EP-00301439.

XX 25-FEB-1999; 99US-0121825P.

XX 05-MAR-1999; 99US-0123180P.

XX 09-MAR-1999; 99US-0123548P.

XX 23-MAR-1999; 99US-0125788P.

XX 25-MAR-1999; 99US-0126284P.

XX 29-MAR-1999; 99US-0126785P.

XX 01-APR-1999; 99US-0127462P.

XX 06-APR-1999; 99US-0128234P.

XX 16-APR-1999; 99US-0128714P.

XX 19-APR-1999; 99US-0129845P.

XX 21-APR-1999; 99US-0130077P.

XX 23-APR-1999; 99US-0130449P.

XX 28-APR-1999; 99US-0130891P.

XX 30-APR-1999; 99US-0131449P.

XX 30-APR-1999; 99US-0132048P.

XX 04-MAY-1999; 99US-0132484P.

XX 05-MAY-1999; 99US-0132485P.

XX 06-MAY-1999; 99US-0132486P.

XX 07-MAY-1999; 99US-0132487P.

XX 11-MAY-1999; 99US-0132863P.

XX 14-MAY-1999; 99US-0134218P.

XX 14-MAY-1999; 99US-0134219P.

XX 14-MAY-1999; 99US-0134221P.

XX 18-MAY-1999; 99US-0134370P.

XX 19-MAY-1999; 99US-0134768P.

XX 20-MAY-1999; 99US-0134941P.

XX 21-MAY-1999; 99US-0135124P.

XX 24-MAY-1999; 99US-0135353P.

XX 24-MAY-1999; 99US-0135629P.





```
PR 28-OCT-1999; 99US-0161993P.
PR 29-OCT-1999; 99US-0162142P.

Query Match      100.0%; Score 75; DB 3; Length 118;
Best Local Similarity 100.0%; Pred. No. 2.7e-06;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      1 LPFAILRLDLAGRL 15
      |||||
Db      46 LPFAILRLDLAGRL 60

RESULT 13
AAG26467
ID AAG26467 standard; protein; 135 AA.
XX AC AAG26467;
XX DT 17-OCT-2000 (first entry)
XX DE Arabidopsis thaliana protein fragment SEQ ID NO: 30933.
XX KW Protein identification; signal transduction pathway; metabolic pathway;
KW hybridisation assay; genetic mapping; gene expression control; promoter;
XX OS Arabidopsis thaliana.
XX FN EP1033405-A2.
XX PD 06-SEP-2000.
XX PF 25-FEB-2000; 2000EP-00301439.
XX PR 25-FEB-1999; 99US-0121825P.
XX PR 05-MAR-1999; 99US-0123180P.
XX PR 09-MAR-1999; 99US-0123548P.
XX PR 23-MAR-1999; 99US-0125788P.
XX PR 25-MAR-1999; 99US-0126264P.
XX PR 29-MAR-1999; 99US-0126785P.
XX PR 01-APR-1999; 99US-0127462P.
XX PR 06-APR-1999; 99US-0128234P.
XX PR 08-APR-1999; 99US-0128714P.
XX PR 16-APR-1999; 99US-0129845P.
XX PR 19-APR-1999; 99US-0130077P.
XX PR 21-APR-1999; 99US-0130449P.
XX PR 23-APR-1999; 99US-0130510P.
XX PR 23-APR-1999; 99US-0130891P.
XX PR 28-APR-1999; 99US-0131449P.
XX PR 30-APR-1999; 99US-0132048P.
XX PR 30-APR-1999; 99US-0132407P.
XX PR 04-MAY-1999; 99US-0132484P.
XX PR 05-MAY-1999; 99US-0132485P.
XX PR 06-MAY-1999; 99US-0132486P.
XX PR 06-MAY-1999; 99US-0132487P.
XX PR 07-MAY-1999; 99US-0132863P.
XX PR 11-MAY-1999; 99US-0134256P.
XX PR 14-MAY-1999; 99US-0134218P.
XX PR 14-MAY-1999; 99US-0134219P.
XX PR 14-MAY-1999; 99US-0134221P.
XX PR 14-MAY-1999; 99US-0134370P.
XX PR 18-MAY-1999; 99US-0134768P.
XX PR 19-MAY-1999; 99US-0134941P.
XX PR 20-MAY-1999; 99US-0135124P.
XX PR 21-MAY-1999; 99US-0135353P.
XX PR 24-MAY-1999; 99US-0135629P.
XX PR 25-MAY-1999; 99US-0136021P.
XX PR 27-MAY-1999; 99US-0136392P.
XX PR 28-MAY-1999; 99US-0136782P.
XX PR 01-JUN-1999; 99US-0137228P.
XX PR 03-JUN-1999; 99US-0137528P.
XX PR 04-JUN-1999; 99US-0137502P.
XX PR 07-JUN-1999; 99US-0137724P.

108-JUN-1999; 99US-0138094P.
10-JUN-1999; 99US-0138540P.
10-JUN-1999; 99US-0138847P.
14-JUN-1999; 99US-0139119P.
16-JUN-1999; 99US-0139452P.
16-JUN-1999; 99US-0139453P.
17-JUN-1999; 99US-0139492P.
18-JUN-1999; 99US-0139454P.
18-JUN-1999; 99US-0139455P.
18-JUN-1999; 99US-0139456P.
18-JUN-1999; 99US-0139457P.
18-JUN-1999; 99US-0139458P.
18-JUN-1999; 99US-0139459P.
18-JUN-1999; 99US-0139460P.
18-JUN-1999; 99US-0139461P.
18-JUN-1999; 99US-0139462P.
18-JUN-1999; 99US-0139463P.
18-JUN-1999; 99US-0139750P.
21-JUN-1999; 99US-0139763P.
21-JUN-1999; 99US-0139817P.
22-JUN-1999; 99US-0139899P.
23-JUN-1999; 99US-0140353P.
23-JUN-1999; 99US-0140354P.
24-JUN-1999; 99US-0140695P.
28-JUN-1999; 99US-0140823P.
29-JUN-1999; 99US-0140991P.
30-JUN-1999; 99US-0141287P.
01-JUL-1999; 99US-0141842P.
01-JUL-1999; 99US-0142154P.
02-JUL-1999; 99US-0142055P.
06-JUL-1999; 99US-0142390P.
08-JUL-1999; 99US-0142803P.
09-JUL-1999; 99US-0142920P.
12-JUL-1999; 99US-0142977P.
13-JUL-1999; 99US-0143542P.
14-JUL-1999; 99US-0143624P.
15-JUL-1999; 99US-0144005P.
16-JUL-1999; 99US-0144085P.
16-JUL-1999; 99US-0144086P.
19-JUL-1999; 99US-0144325P.
19-JUL-1999; 99US-0144331P.
19-JUL-1999; 99US-0144332P.
19-JUL-1999; 99US-0144333P.
19-JUL-1999; 99US-0144334P.
19-JUL-1999; 99US-0144335P.
20-JUL-1999; 99US-0144352P.
20-JUL-1999; 99US-0144632P.
20-JUL-1999; 99US-0144884P.
21-JUL-1999; 99US-0144814P.
21-JUL-1999; 99US-0145086P.
21-JUL-1999; 99US-0145088P.
22-JUL-1999; 99US-0145085P.
22-JUL-1999; 99US-0145087P.
22-JUL-1999; 99US-0145089P.
22-JUL-1999; 99US-0145192P.
23-JUL-1999; 99US-0145145P.
23-JUL-1999; 99US-0145218P.
23-JUL-1999; 99US-0145224P.
26-JUL-1999; 99US-0145276P.
27-JUL-1999; 99US-0145913P.
27-JUL-1999; 99US-0145918P.
27-JUL-1999; 99US-0145919P.
28-JUL-1999; 99US-0145951P.
02-AUG-1999; 99US-0146386P.
02-AUG-1999; 99US-0146388P.
02-AUG-1999; 99US-0146389P.
03-AUG-1999; 99US-0147038P.
04-AUG-1999; 99US-0147204P.
04-AUG-1999; 99US-0147302P.
05-AUG-1999; 99US-0147192P.
05-AUG-1999; 99US-0147260P.
06-AUG-1999; 99US-0147303P.
06-AUG-1999; 99US-0147416P.
```

PR 09-AUG-1999; 99US-0147493P.  
PR 09-AUG-1999; 99US-0147935P.  
PR 10-AUG-1999; 99US-0148171P.  
PR 11-AUG-1999; 99US-0148311P.  
PR 12-AUG-1999; 99US-0148341P.  
PR 13-AUG-1999; 99US-0148565P.  
PR 13-AUG-1999; 99US-0148684P.  
PR 16-AUG-1999; 99US-0149368P.  
PR 17-AUG-1999; 99US-0149175P.  
PR 18-AUG-1999; 99US-0149426P.  
PR 20-AUG-1999; 99US-0149722P.  
PR 20-AUG-1999; 99US-0149723P.  
PR 20-AUG-1999; 99US-0149929P.  
PR 23-AUG-1999; 99US-0149902P.  
PR 23-AUG-1999; 99US-0149930P.  
PR 25-AUG-1999; 99US-0150566P.  
PR 26-AUG-1999; 99US-0150884P.  
PR 27-AUG-1999; 99US-0151065P.  
PR 27-AUG-1999; 99US-0151066P.  
PR 27-AUG-1999; 99US-0151068P.  
PR 30-AUG-1999; 99US-0151303P.  
PR 31-AUG-1999; 99US-0151438P.  
PR 01-SEP-1999; 99US-0151930P.  
PR 07-SEP-1999; 99US-0152363P.  
PR 10-SEP-1999; 99US-0153070P.  
PR 13-SEP-1999; 99US-0153758P.  
PR 15-SEP-1999; 99US-0154018P.  
PR 16-SEP-1999; 99US-0154039P.  
PR 20-SEP-1999; 99US-0154779P.  
PR 22-SEP-1999; 99US-0155139P.  
PR 23-SEP-1999; 99US-0155486P.  
PR 24-SEP-1999; 99US-0155659P.  
PR 28-SEP-1999; 99US-0156458P.  
PR 29-SEP-1999; 99US-0156596P.  
PR 04-OCT-1999; 99US-0157117P.  
PR 05-OCT-1999; 99US-0157753P.  
PR 06-OCT-1999; 99US-0157865P.  
PR 07-OCT-1999; 99US-0158029P.  
PR 08-OCT-1999; 99US-0158232P.  
PR 12-OCT-1999; 99US-0158369P.  
PR 13-OCT-1999; 99US-0159293P.  
PR 13-OCT-1999; 99US-0159294P.  
PR 13-OCT-1999; 99US-0159295P.  
PR 14-OCT-1999; 99US-0159329P.  
PR 14-OCT-1999; 99US-0159330P.  
PR 14-OCT-1999; 99US-0159331P.  
PR 14-OCT-1999; 99US-0159637P.  
PR 14-OCT-1999; 99US-0159638P.  
PR 18-OCT-1999; 99US-0159584P.  
PR 21-OCT-1999; 99US-0160741P.  
PR 21-OCT-1999; 99US-0160767P.  
PR 21-OCT-1999; 99US-0160768P.  
PR 21-OCT-1999; 99US-0160770P.  
PR 21-OCT-1999; 99US-0160814P.  
PR 21-OCT-1999; 99US-0160815P.  
PR 22-OCT-1999; 99US-0160980P.  
PR 22-OCT-1999; 99US-0160981P.  
PR 22-OCT-1999; 99US-0160981P.  
PR 22-OCT-1999; 99US-0160982P.  
PR 25-OCT-1999; 99US-0161404P.  
PR 25-OCT-1999; 99US-0161405P.  
PR 25-OCT-1999; 99US-0161406P.  
PR 26-OCT-1999; 99US-0161359P.  
PR 26-OCT-1999; 99US-0161360P.  
PR 26-OCT-1999; 99US-0161361P.  
PR 28-OCT-1999; 99US-0161920P.  
PR 28-OCT-1999; 99US-0161992P.  
PR 28-OCT-1999; 99US-0161993P.  
PR 29-OCT-1999; 99US-0162142P.  
PR 29-OCT-1999; 99US-0162143P.

Query Match 100.0%; Score 75; DB 3; Length 135;  
Best Local Similarity 100.0%; Pred. No. 3.1e-06;  
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 LPFAILRLDLGRDL 15  
Db 63 LPFAILRLDLGRDL 77

## RESULT 14

ADK36534  
ID ADK36534 standard; protein; 148 AA.

XX AC ADK36534;

XX DT 06-MAY-2004 (first entry)

XX DE Novel human polypeptide SeqID8616.

XX KW antiarthritic; antiparkinsonian; neuroprotective; nootropic;  
immunosuppressive; cytostatic; antiparasitic; antiinflammatory;  
antibacterial; antiviral; antifungal; antiparasitic; gene therapy;  
arthritis; Parkinson's; Alzheimer's; autoimmune disease; cancer;  
psoriasis; inflammatory bowel disease; infection; bacteria; virus;  
fungus; parasite; human.

XX OS Homo sapiens.

XX FH Key Location/Qualifiers

FT Misc-difference 1..148

FT /label= OTHER

FT /note= "OTHER= All Xaa's in this sequence are unknown  
amino acids or the site of a stop codon within the DNA  
sequence"

XX PN WO200216439-A2.

XX PD 28-FEB-2002.

XX PF 05-MAR-2001; 2001WO-US004941.

XX PR 07-MAR-2000; 2000US-00519705.

XX PR 19-MAY-2000; 2000US-00574454.

XX PA (HYSE-) HYSEQ INC.

XX PI Tang YT, Liu C, Drmanac RT;

XX WPI; 2002-280918/32.

PT Isolated polynucleotide encoding bone marrow derived polypeptides useful  
for treating, e.g., Parkinson's, Alzheimer's, cancer, arthritis, Crohn's  
disease, and inflammatory bowel disease.

PS Claim 20; SEQ ID NO 8616; 504pp; English.

XX This invention relates to a novel isolated polynucleotide comprising a  
nucleotide sequence selected from one of 1680 sequences, a mature protein  
coding portion of them, an active domain of them and their complementary  
sequences. The invention may be useful for the production of compounds  
with an antiarthritic, antiparkinsonian, neuroprotective, nootropic,  
immunosuppressive, cytostatic, antiparasitic, antiinflammatory,  
antibacterial, antiviral, antifungal or antiparasitic activity. In  
addition, the disclosed sequences may be useful for gene therapy. The  
polypeptides or their antibodies are useful for treating many diseases  
such as arthritis, Parkinson's, Alzheimer's, autoimmune diseases, cancer,  
psoriasis, inflammatory bowel disease and infections caused by bacteria,  
viruses, fungi or parasites. The present sequence is that of a human  
polypeptide of the invention.

SQ Sequence 148 AA;

Query Match 100.0%; Score 75; DB 5; Length 148;

Best Local Similarity 100.0%; Pred. No. 3.5e-06;  
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 LPFAILRLDLGRDL 15

Search completed: April 8, 2005, 10:50:46  
Job time : 63.4 secs

Db 30 LPHALLRLDLGRDL 44

RESULT 15  
ABG15101  
ID ABG15101 standard; protein; 150 AA.  
XX AC ABG15101;  
XX DT 18-FEB-2002 (first entry)  
XX DE Novel human diagnostic protein #15092.  
XX KW Human; chromosome mapping; gene mapping; gene therapy; forensic;  
XX KW food supplement; medical imaging; diagnostic; genetic disorder.  
XX OS Homo sapiens.  
XX PN WO200175067-A2.  
XX PD 11-OCT-2001.  
XX PF 30-MAR-2001; 2001WO-US008631.  
XX PR 31-MAR-2000; 2000US-00540217.  
XX PR 23-AUG-2000; 2000US-00649167.  
XX PA (HYSE-) HYSEQ INC.  
XX PI Drmanac RT, Liu C, Tang YT;  
XX DR WPI: 2001-639362/73.  
XX DR N-PSDB; AAS79288.  
XX PT New isolated polynucleotide and encoded polypeptides, useful in  
XX PT diagnostics, forensics, gene mapping, identification of mutations  
XX PT responsible for genetic disorders or other traits and to assess  
XX PT biodiversity.  
XX PS Claim 20; SEQ ID NO 45460; 103pp; English.  
XX CC The invention relates to isolated polynucleotide (I) and polypeptide (II)  
XX CC sequences. (I) is useful as hybridisation probes, polymerase chain  
XX CC reaction (PCR) primers, oligomers, and for chromosome and gene mapping,  
XX CC and in recombinant production of (II). The polynucleotides are also used  
XX CC in diagnostics as expressed sequence tags for identifying expressed  
XX CC genes. (I) is useful in gene therapy techniques to restore normal  
XX CC activity of (II) or to treat disease states involving (II). (II) is  
XX CC useful for generating antibodies against it, detecting or quantitating a  
XX CC polypeptide in tissue, as molecular weight markers and as a food  
XX CC supplement. (II) and its binding partners are useful in medical imaging  
XX CC of sites expressing (II). (I) and (II) are useful for treating disorders  
XX CC involving aberrant protein expression or biological activity. The  
XX CC polypeptide and polynucleotide sequences have applications in  
XX CC diagnostics, forensics, gene mapping, identification of mutations  
XX CC responsible for genetic disorders or other traits to assess biodiversity  
XX CC and to produce other types of data and products dependent on DNA and  
XX CC amino acid sequences. ABG00010-ABG30377 represent novel human diagnostic  
XX CC amino acid sequences of the invention. Note: The sequence data for this  
XX CC patent did not appear in the printed specification, but was obtained in  
XX CC electronic format directly from WIPO at  
XX CC ftp.wipo.int/pub/published\_pct\_sequences  
XX SQ Sequence 150 AA;

Query Match 100.0%; Score 75; DB 4; Length 150;  
Best Local Similarity 100.0%; Pred. No. 3.6e-06;  
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 LPHALLRLDLGRDL 15  
133 LPHALLRLDLGRDL 147

Db

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: April 8, 2005, 09:58:57 ; Search time 17.5333 Seconds  
(without alignments)  
63.863 Million cell updates/sec

Title: US-09-423-351C-7  
Perfect score: 75  
Sequence: 1 LPHAILRLDLGRDL 15

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 513545 seqs, 74649064 residues

Total number of hits satisfying chosen parameters: 513545

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : Issued Patents AA:\*  
1: /cgn2\_6/ptodata/1/1aa/5A\_COMB.pep:\*  
2: /cgn2\_6/ptodata/1/1aa/5B\_COMB.pep:\*  
3: /cgn2\_6/ptodata/1/1aa/6A\_COMB.pep:\*  
4: /cgn2\_6/ptodata/1/1aa/6B\_COMB.pep:\*  
5: /cgn2\_6/ptodata/1/1aa/PTUS\_COMB.pep:\*  
6: /cgn2\_6/ptodata/1/1aa/backfiles1.pep:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	75	100.0	146	3 US-09-306-446C-5	Sequence 5, Appli
2	75	100.0	374	3 US-08-609-236-6	Sequence 6, Appli
3	75	100.0	374	3 US-09-306-446C-2	Sequence 2, Appli
4	75	100.0	375	2 US-08-494-151-14	Sequence 14, Appl
5	75	100.0	375	4 US-09-976-594-731	Sequence 731, Appl
6	75	100.0	399	4 US-09-949-016-9424	Sequence 9424, Ap
7	75	100.0	404	4 US-09-949-016-11313	Sequence 11313, A
8	73	97.3	239	4 US-09-949-016-10409	Sequence 10409, A
9	73	97.3	375	3 US-09-106-217-16	Sequence 16, Appl
10	73	97.3	376	1 US-08-588-113-2	Sequence 2, Appli
11	73	97.3	376	4 US-09-949-016-6100	Sequence 6100, Ap
12	73	97.3	377	3 US-09-106-217-2	Sequence 2, Appli
13	73	97.3	377	4 US-09-917-254-53	Sequence 53, Appl
14	73	97.3	386	4 US-09-949-016-7721	Sequence 7721, Ap
15	73	97.3	402	4 US-09-949-016-10757	Sequence 10757, A
16	69	92.0	377	4 US-09-248-796A-14109	Sequence 14109, A
17	65	86.7	375	3 US-09-171-337A-8	Sequence 8, Appli
18	65	86.7	375	4 US-09-631-022-8	Sequence 8, Appli
19	64	85.3	375	3 US-09-171-337A-7	Sequence 7, Appli
20	64	85.3	375	4 US-09-631-022-7	Sequence 7, Appli
21	64	85.3	377	4 US-09-919-172-33	Sequence 33, Appl
22	61	81.3	362	4 US-09-949-016-7725	Sequence 7725, Ap
23	61	81.3	371	1 US-08-261-208A-77	Sequence 77, Appl
24	61	81.3	376	4 US-09-538-092-1109	Sequence 1109, Ap
25	60	80.0	376	4 US-09-538-092-1110	Sequence 1110, Ap
26	60	80.0	376	4 US-09-949-016-6656	Sequence 6656, Ap
27	60	80.0	376	4 US-09-949-016-8452	Sequence 8452, Ap

28	73.3	394	4	US-09-949-016-6655	Sequence 6655, Ap
29	73.3	406	4	US-09-949-016-7396	Sequence 7396, Ap
30	66.7	336	4	US-09-248-796A-14108	Sequence 14108, A
31	61.3	670	4	US-09-252-991A-22079	Sequence 22079, A
32	60.0	327	4	US-09-252-991A-29837	Sequence 29837, A
33	56.0	1056	4	US-09-252-991A-24162	Sequence 24162, A
34	54.7	176	4	US-09-902-540-13432	Sequence 13432, A
35	54.7	354	4	US-09-252-991A-18669	Sequence 18669, A
36	54.7	384	4	US-09-538-092-395	Sequence 395, App
37	54.7	925	4	US-09-252-991A-20697	Sequence 20697, A
38	54.7	1121	4	US-08-915-048A-2	Sequence 2, Appli
39	53.3	550	4	US-09-252-991A-21032	Sequence 21032, A
40	53.3	917	4	US-09-902-540-12845	Sequence 12845, A
41	53.3	1180	4	US-09-252-991A-32464	Sequence 32464, A
42	52.0	182	4	US-09-583-110-4365	Sequence 4365, Ap
43	52.0	187	4	US-09-107-433-2893	Sequence 2893, Ap
44	52.0	271	4	US-09-252-991A-30514	Sequence 30514, A
45	52.0	425	3	US-09-036-315-5	Sequence 5, Appli

## ALIGNMENTS

RESULT 1  
US-09-306-446C-5  
; Sequence 5, Application US/09306446C  
; Patent No. 6372959  
; GENERAL INFORMATION:  
; APPLICANT: KIM, Dong Soo  
; APPLICANT: KIM, Chul Geun  
; APPLICANT: NAM, Yoon Kwon  
; APPLICANT: NOH, Jae Koo  
; APPLICANT: CHO, Kyoun Nam  
; TITLE OF INVENTION: EXPRESSION VECTOR OF MUD LOACH GROWTH HORMONE GENE  
; FILE REFERENCE: P06344US/0/BAS  
; CURRENT APPLICATION NUMBER: US/09/306,446C  
; CURRENT FILING DATE: 1999-05-06  
; PRIOR APPLICATION NUMBER: KR 98/20255  
; PRIOR FILING DATE: 1998-06-01  
; NUMBER OF SEQ ID NOS: 33  
; SOFTWARE: Patentin Ver. 2.0  
; SEQ ID NO 5  
; LENGTH: 146  
; TYPE: PRT  
; ORGANISM: Misgurnus mizolepus  
US-09-306-446C-5

Query Match 100.0%; Score 75; DB 3; Length 146;  
Best Local Similarity 100.0%; Pred. No. 2.5e-06;  
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 LPHAILRLDLGRDL 15  
DB 50 LPHAILRLDLGRDL 64  
RESULT 2  
US-08-609-236-6  
; Sequence 6, Application US/08609236  
; Patent No. 6087398  
; GENERAL INFORMATION:  
; APPLICANT: Steven R. Goodman  
; TITLE OF INVENTION: No. 6087398el Sickle Cell Anemia Treatment  
; NUMBER OF SEQUENCES: 6  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: McGregor & Adler, LLP  
; STREET: 8011 Candle Lane  
; CITY: Houston  
; STATE: Texas  
; COUNTRY: USA  
; ZIP: 77071  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk

```
; COMPUTER: Apple Macintosh
; OPERATING SYSTEM: Macintosh
; SOFTWARE: Microsoft Word for Macintosh
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/609,236
; FILING DATE: March 1, 1996
; CLASSIFICATION: 514
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 60/002,288
; FILING DATE: August 14, 1995
; CLASSIFICATION: 514
; ATTORNEY/AGENT INFORMATION:
; NAME: Benjamin Aaron Adler, Ph.D.
; REGISTRATION NUMBER: 35,423
; REFERENCE/DOCKET NUMBER: D5807
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 713-777-2321
; TELEFAX: 713-777-6908
; INFORMATION FOR SEQ ID NO: 6:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 374
; TYPE: Amino acid
; STRANDEDNESS:
; TOPOLOGY: Linear
; MOLECULE TYPE: Protein
; DESCRIPTION: Protein
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
; FRAGMENT TYPE:
; ORIGINAL SOURCE:
; STRAIN:
; INDIVIDUAL ISOLATE:
; DEVELOPMENTAL STAGE:
; TISSUE TYPE:
; CELL TYPE:
; CELL LINE:
;
US-08-609-236-5

Query Match 100.0%; Score 75; DB 3; Length 374;
Best Local Similarity 100.0%; Pred. No. 7.7e-06;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 LPHAILRLDLGRDL 15
Db 170 LPHAILRLDLGRDL 184

RESULT 3
US-09-306-446C-2
; Sequence 2, Application US/09306446C
; Patent No. 6372959
; GENERAL INFORMATION:
; APPLICANT: KIM, Dong Soo
; APPLICANT: KIM, Chul Geun
; APPLICANT: NAM, Yoon Kwon
; APPLICANT: NOH, Jae Koo
; APPLICANT: CHO, Kyou Nam
; TITLE OF INVENTION: EXPRESSION VECTOR OF MUD LOACH GROWTH HORMONE GENE
; FILE REFERENCE: P06344US0/BAS
; CURRENT APPLICATION NUMBER: US/09/306,446C
; CURRENT FILING DATE: 1999-05-06
; PRIOR APPLICATION NUMBER: KR 98/20255
; PRIOR FILING DATE: 1998-06-01
; NUMBER OF SEQ ID NOS: 33
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 2
; LENGTH: 374
; TYPE: PRT
; ORGANISM: Misgurnus mizolepus
;
US-09-306-446C-2

Query Match 100.0%; Score 75; DB 3; Length 374;
Best Local Similarity 100.0%; Pred. No. 7.7e-06;
```

```
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 LPHAILRLDLGRDL 15
Db 171 LPHAILRLDLGRDL 185

RESULT 4
US-08-494-151-14
; Sequence 14, Application US/08494151
; Patent No. 5840528
; GENERAL INFORMATION:
; APPLICANT: Van Ooyen, Albert Johannes Joseph
; TITLE OF INVENTION: Transformation of Phaffia rhodozyma
; NUMBER OF SEQUENCES: 14
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Morrison & Foerster
; STREET: 2000 Pennsylvania Ave. N.W., Suite 5500
; CITY: Washington, D.C.
; COUNTRY: USA
; ZIP: 20006-1812
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA: US/08/494,151
; APPLICATION NUMBER: US/08/494,151
; FILING DATE: 23-JUN-1995
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Murashige, Kate H.
; REGISTRATION NUMBER: 29,959
; REFERENCE/DOCKET NUMBER: 24615-20039.01
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202) 887-1500
; TELEFAX: (202) 887-0763
; TELEX: 90-4030
; INFORMATION FOR SEQ ID NO: 14:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 375 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
;
US-08-494-151-14

Query Match 100.0%; Score 75; DB 2; Length 375;
Best Local Similarity 100.0%; Pred. No. 7.7e-06;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 LPHAILRLDLGRDL 15
Db 171 LPHAILRLDLGRDL 185

RESULT 5
US-09-976-594-731
; Sequence 731, Application US/09976594
; Patent No. 6673549
; GENERAL INFORMATION:
; APPLICANT: Furness, Michael
; APPLICANT: Buchbinder, Jenny
; TITLE OF INVENTION: GENES EXPRESSED IN C3A LIVER CELL CULTURES TREATED WITH STEROIDS
; FILE REFERENCE: PA-0041 US
; CURRENT APPLICATION NUMBER: US/09/976,594
; CURRENT FILING DATE: 2001-10-12
; PRIOR APPLICATION NUMBER: 60/240,409
; PRIOR FILING DATE: 2000-10-12
; NUMBER OF SEQ ID NOS: 1143
; SOFTWARE: PERL Program
; SEQ ID NO 731
; LENGTH: 375
; TYPE: PRT
```

```
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; OTHER INFORMATION: Incyte ID No. 6673549 1837317CD1
US-09-976-594-731

Query Match      100.0%; Score 75; DB 4; Length 375;
Best Local Similarity 100.0%; Pred. No. 7.7e-06;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      1 LPHAILRLDLGRDL 15
Db      171 LPHAILRLDLGRDL 185

RESULT 6
US-09-949-016-9424
; Sequence 9424, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 9424
; LENGTH: 399
; TYPE: PRT
; ORGANISM: Human
US-09-949-016-9424

Query Match      100.0%; Score 75; DB 4; Length 399;
Best Local Similarity 100.0%; Pred. No. 8.3e-06;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      1 LPHAILRLDLGRDL 15
Db      195 LPHAILRLDLGRDL 209

RESULT 7
US-09-949-016-11313
; Sequence 11313, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 11313
; LENGTH: 404
; TYPE: PRT
; ORGANISM: Human
US-09-949-016-11313
```

```
Query Match      100.0%; Score 75; DB 4; Length 404;
Best Local Similarity 100.0%; Pred. No. 8.4e-06;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      1 LPHAILRLDLGRDL 15
Db      200 LPHAILRLDLGRDL 214

RESULT 8
US-09-949-016-10409
; Sequence 10409, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 10409
; LENGTH: 239
; TYPE: PRT
; ORGANISM: Human
US-09-949-016-10409

Query Match      97.3%; Score 73; DB 4; Length 239;
Best Local Similarity 93.3%; Pred. No. 1.1e-05;
Matches 14; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy      1 LPHAILRLDLGRDL 15
Db      35 LPHAILRLDLGRDL 49

RESULT 9
US-09-106-217-16
; Sequence 16, Application US/09106217
; Patent No. 6063576
; GENERAL INFORMATION:
; APPLICANT: Keating, Mark T.
; APPLICANT: Olson, Timothy M.
; TITLE OF INVENTION: Actin Mutations in Dilated
; TITLE OF INVENTION: Cardiomyopathy, a Heritable Form of Heart Failure
; NUMBER OF SEQUENCES: 18
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Rothwell, Figg, Ernst & Kurz, P.C.
; STREET: 555 Thirteenth Street, N.W., Suite 701 East
; STREET: Tower
; CITY: Washington
; STATE: DC
; COUNTRY: U.S.A.
; ZIP: 20004
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/106,217
; FILING DATE:
; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
; NAME: Saxe, Stephen A.
```

REGISTRATION NUMBER: 38,609  
REFERENCE/DOCKET NUMBER: 2323-125  
TELEPHONE: 202-783-6040  
TELEFAX: 202-783-6031  
INFORMATION FOR SEQ ID NO: 16:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 375 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-09-106-217-16

Query Match 97.3%; Score 73; DB 3; Length 375;  
Best Local Similarity 93.3%; Pred. No. 1.8e-05;  
Matches 14; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 LPHAILRLDLGRDL 15  
||||:|||||  
Db 171 LPHAIMRLDLGRDL 185

RESULT 10  
US-08-588-113-2  
Sequence 2, Application US/08588113  
Patent No. 5710003  
GENERAL INFORMATION:  
APPLICANT: McHugh, Kirk M.  
TITLE OF INVENTION: DIAGNOSTIC TEST FOR DETERMINING  
MALIGNANCY OF SMOOTH MUSCLE TUMORS  
NUMBER OF SEQUENCES: 4  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Woodcock Washburn Kurtz Mackiewicz &  
ADDRESSEE: No 571003ris  
STREET: One Liberty Place - 46th Floor  
CITY: Philadelphia  
STATE: PA  
COUNTRY: USA  
ZIP: 19103  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
FILING DATE:  
CLASSIFICATION: 436  
ATTORNEY/AGENT INFORMATION:  
NAME: Ralph, Rebecca L.  
REGISTRATION NUMBER: 35,152  
REFERENCE/DOCKET NUMBER: TJU-1652  
TELEPHONE: 215-568-3100  
TELEFAX: 215-568-3439  
INFORMATION FOR SEQ ID NO: 2:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 376 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-588-113-2

Query Match 97.3%; Score 73; DB 1; Length 376;  
Best Local Similarity 93.3%; Pred. No. 1.8e-05;  
Matches 14; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 LPHAILRLDLGRDL 15  
||||:|||||  
Db 172 LPHAIMRLDLGRDL 186

RESULT 11

US-09-949-016-6100  
Sequence 6100, Application US/09949016  
Patent No. 6812339  
GENERAL INFORMATION:  
APPLICANT: VENTER, J. Craig et al.  
TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED  
WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF  
FILE REFERENCE: CL001307  
CURRENT APPLICATION NUMBER: US/09/949,016  
CURRENT FILING DATE: 2000-04-14  
PRIOR APPLICATION NUMBER: 60/241,755  
PRIOR FILING DATE: 2000-10-20  
PRIOR APPLICATION NUMBER: 60/237,768  
PRIOR FILING DATE: 2000-10-03  
PRIOR APPLICATION NUMBER: 60/231,498  
PRIOR FILING DATE: 2000-09-08  
NUMBER OF SEQ ID NOS: 207012  
SOFTWARE: FastSeq for Windows Version 4.0  
SEQ ID NO 6100  
LENGTH: 376  
TYPE: PRT  
ORGANISM: Human  
US-09-949-016-6100

Query Match 97.3%; Score 73; DB 4; Length 376;  
Best Local Similarity 93.3%; Pred. No. 1.8e-05;  
Matches 14; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 LPHAILRLDLGRDL 15  
||||:|||||  
Db 172 LPHAIMRLDLGRDL 186

RESULT 12  
US-09-106-217-2  
Sequence 2, Application US/09106217  
Patent No. 6063576  
GENERAL INFORMATION:  
APPLICANT: Keating, Mark T.  
APPLICANT: Olson, Timothy M.  
TITLE OF INVENTION: Actin Mutations in Dilated  
Cardiomyopathy, a Heritable Form of Heart Failure  
NUMBER OF SEQUENCES: 18  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Rotwell, Figg, Ernst & Kurz, P.C.  
STREET: 555 Thirteenth Street, N.W., Suite 701 East  
STREET: Tower  
CITY: Washington  
STATE: DC  
COUNTRY: U.S.A.  
ZIP: 20004  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
FILING DATE:  
CLASSIFICATION:  
ATTORNEY/AGENT INFORMATION:  
NAME: Saxe, Stephen A.  
REGISTRATION NUMBER: 38,609  
REFERENCE/DOCKET NUMBER: 2323-125  
TELEPHONE: 202-783-6040  
TELEFAX: 202-783-6031  
INFORMATION FOR SEQ ID NO: 2:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 377 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein



US-09-106-217-2

Query Match 97.3%; Score 73; DB 3; Length 377;  
Best Local Similarity 93.3%; Pred. No. 1.8e-05;  
Matches 14; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 LPHAILRLDLAQRDL 15  
Db 173 LPHAIMRLDLAQRDL 187

RESULT 13

US-09-917-254-53  
; Sequence 53, Application US/09917254  
; Patent No. 6703204  
; GENERAL INFORMATION:  
; APPLICANT: Mutter, George  
; APPLICANT: Baak, Jan  
; TITLE OF INVENTION: Prognostic Classification of Breast Cancer  
; FILE REFERENCE: B0801/7224(JRV)  
; CURRENT APPLICATION NUMBER: US/09/917,254  
; CURRENT FILING DATE: 2001-07-27  
; PRIOR APPLICATION NUMBER: US 60/222,093  
; PRIOR FILING DATE: 2000-07-28  
; NUMBER OF SEQ ID NOS: 102  
; SOFTWARE: PatentIn version 3.0  
; SEQ ID NO 53  
; LENGTH: 377  
; TYPE: PRT  
; ORGANISM: Homo Sapiens  
US-09-917-254-53

Query Match 97.3%; Score 73; DB 4; Length 377;  
Best Local Similarity 93.3%; Pred. No. 1.8e-05;  
Matches 14; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 LPHAILRLDLAQRDL 15  
Db 173 LPHAIMRLDLAQRDL 187

RESULT 14

US-09-949-016-7721  
; Sequence 7721, Application US/09949016  
; Patent No. 6812339  
; GENERAL INFORMATION:  
; APPLICANT: VENTER, J. Craig et al.  
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED  
; FILE REFERENCE: CL001307  
; CURRENT APPLICATION NUMBER: US/09/949,016  
; CURRENT FILING DATE: 2000-04-14  
; PRIOR APPLICATION NUMBER: 60/241,755  
; PRIOR FILING DATE: 2000-10-20  
; PRIOR APPLICATION NUMBER: 60/237,768  
; PRIOR FILING DATE: 2000-10-03  
; PRIOR APPLICATION NUMBER: 60/231,498  
; PRIOR FILING DATE: 2000-09-08  
; NUMBER OF SEQ ID NOS: 207012  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 7721  
; LENGTH: 386  
; TYPE: PRT  
; ORGANISM: Human  
US-09-949-016-7721

Query Match 97.3%; Score 73; DB 4; Length 386;  
Best Local Similarity 93.3%; Pred. No. 1.9e-05;  
Matches 14; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 LPHAILRLDLAQRDL 15  
Db 182 LPHAIMRLDLAQRDL 196

RESULT 15

US-09-949-016-10757  
; Sequence 10757, Application US/09949016  
; Patent No. 6812339  
; GENERAL INFORMATION:  
; APPLICANT: VENTER, J. Craig et al.  
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED  
; FILE REFERENCE: CL001307  
; CURRENT APPLICATION NUMBER: US/09/949,016  
; CURRENT FILING DATE: 2000-04-14  
; PRIOR APPLICATION NUMBER: 60/241,755  
; PRIOR FILING DATE: 2000-10-20  
; PRIOR APPLICATION NUMBER: 60/237,768  
; PRIOR FILING DATE: 2000-10-03  
; PRIOR APPLICATION NUMBER: 60/231,498  
; PRIOR FILING DATE: 2000-09-08  
; NUMBER OF SEQ ID NOS: 207012  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 10757  
; LENGTH: 402  
; TYPE: PRT  
; ORGANISM: Human  
US-09-949-016-10757

Query Match 97.3%; Score 73; DB 4; Length 402;  
Best Local Similarity 93.3%; Pred. No. 2e-05;  
Matches 14; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 LPHAILRLDLAQRDL 15  
Db 198 LPHAIMRLDLAQRDL 212

Search completed: April 8, 2005, 12:07:36  
Job time : 17.5333 secs

**This Page Blank (uspto)**

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: April 8, 2005, 10:53:18 ; Search time 42.6667 Seconds  
(without alignments)  
116.718 Million cell updates/sec

Title: US-09-423-351C-7

Perfect score: 75

Sequence: 1 LPHAILRLDLGRDL 15

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1418010 seqs, 33197259 residues

Total number of hits satisfying chosen parameters: 1418010

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications AA:\*

- 1: /cgn2\_6/ptodata/1/pubpaa/US07\_PUBCOMB.pep.\*
- 2: /cgn2\_6/ptodata/1/pubpaa/PCT\_NEW\_PUB.pep.\*
- 3: /cgn2\_6/ptodata/1/pubpaa/US06\_NEW\_PUB.pep.\*
- 4: /cgn2\_6/ptodata/1/pubpaa/US06\_PUBCOMB.pep.\*
- 5: /cgn2\_6/ptodata/1/pubpaa/US07\_NEW\_PUB.pep.\*
- 6: /cgn2\_6/ptodata/1/pubpaa/PCTUS\_PUBCOMB.pep.\*
- 7: /cgn2\_6/ptodata/1/pubpaa/US08\_NEW\_PUB.pep.\*
- 8: /cgn2\_6/ptodata/1/pubpaa/US08\_PUBCOMB.pep.\*
- 9: /cgn2\_6/ptodata/1/pubpaa/US09A\_PUBCOMB.pep.\*
- 10: /cgn2\_6/ptodata/1/pubpaa/US09B\_PUBCOMB.pep.\*
- 11: /cgn2\_6/ptodata/1/pubpaa/US09C\_PUBCOMB.pep.\*
- 12: /cgn2\_6/ptodata/1/pubpaa/US09\_NEW\_PUB.pep.\*
- 13: /cgn2\_6/ptodata/1/pubpaa/US10A\_PUBCOMB.pep.\*
- 14: /cgn2\_6/ptodata/1/pubpaa/US10B\_PUBCOMB.pep.\*
- 15: /cgn2\_6/ptodata/1/pubpaa/US10C\_PUBCOMB.pep.\*
- 16: /cgn2\_6/ptodata/1/pubpaa/US10D\_PUBCOMB.pep.\*
- 17: /cgn2\_6/ptodata/1/pubpaa/US10\_NEW\_PUB.pep.\*
- 18: /cgn2\_6/ptodata/1/pubpaa/US11\_NEW\_PUB.pep.\*
- 19: /cgn2\_6/ptodata/1/pubpaa/US60\_NEW\_PUB.pep.\*
- 20: /cgn2\_6/ptodata/1/pubpaa/US60\_PUBCOMB.pep.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	75	100.0	44	15	US-10-424-599-214884
2	75	100.0	97	9	US-09-864-761-33414
3	75	100.0	156	15	US-10-424-599-175571
4	75	100.0	168	15	US-10-264-049-4044
5	75	100.0	180	16	US-10-437-963-169247
6	75	100.0	219	14	US-10-002-631C-56
7	75	100.0	230	16	US-10-767-701-42770
8	75	100.0	236	15	US-10-424-599-206389
9	75	100.0	238	16	US-10-437-963-162706
10	75	100.0	249	16	US-10-767-701-43575
11	75	100.0	250	15	US-10-425-114-44494
12	75	100.0	254	15	US-10-425-114-69353
13	75	100.0	261	15	US-10-424-599-177530

14	75	100.0	276	15	US-10-425-114-52089	Sequence 52089, A
15	75	100.0	288	15	US-10-425-114-57500	Sequence 57500, A
16	75	100.0	299	15	US-10-424-599-278431	Sequence 278431, A
17	75	100.0	313	15	US-10-425-114-62158	Sequence 62158, A
18	75	100.0	331	15	US-10-425-114-59172	Sequence 59172, A
19	75	100.0	347	15	US-10-424-599-283333	Sequence 283333, A
20	75	100.0	371	16	US-10-322-281-68	Sequence 68, Appl
21	75	100.0	375	14	US-10-205-194-93	Sequence 93, Appl
22	75	100.0	375	14	US-10-316-253-88	Sequence 88, Appl
23	75	100.0	375	15	US-10-305-331-94	Sequence 94, Appl
24	75	100.0	375	15	US-10-260-708-82	Sequence 82, Appl
25	75	100.0	376	16	US-10-437-963-179917	Sequence 179917, A
26	75	100.0	376	16	US-10-437-963-202420	Sequence 202420, A
27	75	100.0	376	16	US-10-767-701-45677	Sequence 45677, A
28	75	100.0	376	16	US-10-767-701-46090	Sequence 46090, A
29	75	100.0	377	14	US-10-338-777-52	Sequence 52, Appl
30	75	100.0	377	15	US-10-424-599-162009	Sequence 162009, A
31	75	100.0	377	15	US-10-424-599-162685	Sequence 162685, A
32	75	100.0	377	15	US-10-424-599-162686	Sequence 162686, A
33	75	100.0	377	15	US-10-424-599-222816	Sequence 222816, A
34	75	100.0	377	15	US-10-424-599-223492	Sequence 223492, A
35	75	100.0	377	15	US-10-424-599-283336	Sequence 283336, A
36	75	100.0	377	15	US-10-425-114-66124	Sequence 66124, A
37	75	100.0	377	16	US-10-437-963-121952	Sequence 121952, A
38	75	100.0	377	16	US-10-437-963-148877	Sequence 148877, A
39	75	100.0	377	16	US-10-437-963-198295	Sequence 198295, A
40	75	100.0	377	16	US-10-767-701-47239	Sequence 47239, A
41	75	100.0	378	15	US-10-425-114-36824	Sequence 36824, A
42	75	100.0	378	15	US-10-425-114-42317	Sequence 42317, A
43	75	100.0	378	15	US-10-425-114-52458	Sequence 52458, A
44	75	100.0	378	15	US-10-425-114-59265	Sequence 59265, A
45	75	100.0	378	15	US-10-425-114-59967	Sequence 59967, A

#### ALIGNMENTS

##### RESULT 1

US-10-424-599-214884  
; Sequence 214884, Application US/10424599  
; Publication No. US20040031072A1  
; GENERAL INFORMATION:  
; APPLICANT: La Rosa Thomas J  
; APPLICANT: Kovalic David K  
; APPLICANT: Zhou Yihua  
; APPLICANT: Cao Yongwei  
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With  
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement  
; FILE REFERENCE: 38-21(53223)B  
; CURRENT APPLICATION NUMBER: US/10/424,599  
; CURRENT FILING DATE: 2003-04-28  
; NUMBER OF SEQ ID NOS: 285684  
; SEQ ID NO 214884  
; LENGTH: 44  
; TYPE: PRT  
; ORGANISM: Glycine max  
; FEATURE:  
; OTHER INFORMATION: Clone ID: PAT\_MRT3847\_36066C.1.pep  
US-10-424-599-214884

Query Match 100.0%; Score 75; DB 15; Length 44;  
Best Local Similarity 100.0%; Pred. No. 8.8e-06;  
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 LPHAILRLDLGRDL 15

Db 27 LPHAILRLDLGRDL 41

##### RESULT 2

US-09-864-761-33414  
; Sequence 33414, Application US/09864761  
; Patent No. US20020048763A1

GENERAL INFORMATION:  
APPLICANT: Penn, Sharron G.  
APPLICANT: Rank, David R.  
APPLICANT: Hanzel, David K.  
APPLICANT: Chen, Wensheng  
TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR  
TITLE OF INVENTION: GENE EXPRESSION ANALYSIS BY MICROARRAY  
FILE REFERENCE: Aeomica-X-1  
CURRENT APPLICATION NUMBER: US/09/864,761  
CURRENT FILING DATE: 2001-05-23  
PRIOR APPLICATION NUMBER: US 60/180,312  
PRIOR FILING DATE: 2000-02-04  
PRIOR APPLICATION NUMBER: US 60/207,456  
PRIOR FILING DATE: 2000-05-26  
PRIOR APPLICATION NUMBER: US 09/632,366  
PRIOR FILING DATE: 2000-08-03  
PRIOR APPLICATION NUMBER: GB 24263.6  
PRIOR FILING DATE: 2000-10-04  
PRIOR APPLICATION NUMBER: US 60/236,359  
PRIOR FILING DATE: 2000-09-27  
PRIOR APPLICATION NUMBER: PCT/US01/00666  
PRIOR FILING DATE: 2001-01-30  
PRIOR APPLICATION NUMBER: PCT/US01/00667  
PRIOR FILING DATE: 2001-01-30  
PRIOR APPLICATION NUMBER: PCT/US01/00664  
PRIOR FILING DATE: 2001-01-30  
PRIOR APPLICATION NUMBER: PCT/US01/00669  
PRIOR FILING DATE: 2001-01-30  
PRIOR APPLICATION NUMBER: PCT/US01/00665  
PRIOR FILING DATE: 2001-01-30  
PRIOR APPLICATION NUMBER: PCT/US01/00668  
PRIOR FILING DATE: 2001-01-30  
PRIOR APPLICATION NUMBER: PCT/US01/00663  
PRIOR FILING DATE: 2001-01-30  
PRIOR APPLICATION NUMBER: PCT/US01/00662  
PRIOR FILING DATE: 2001-01-30  
PRIOR APPLICATION NUMBER: PCT/US01/00661  
PRIOR FILING DATE: 2001-01-30  
PRIOR APPLICATION NUMBER: PCT/US01/00670  
PRIOR FILING DATE: 2001-01-30  
PRIOR APPLICATION NUMBER: US 60/234,687  
PRIOR FILING DATE: 2000-09-21  
PRIOR APPLICATION NUMBER: US 09/608,408  
PRIOR FILING DATE: 2000-06-30  
PRIOR APPLICATION NUMBER: US 09/774,203  
PRIOR FILING DATE: 2001-01-29  
NUMBER OF SEQ ID NOS: 49117  
SOFTWARE: Annonax Sequence Listing Engine vers. 1.1  
SEQ ID NO 33414  
LENGTH: 97  
TYPE: PRT  
ORGANISM: Homo sapiens  
FEATURE:  
OTHER INFORMATION: MAP TO AC006483.2  
OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 18  
OTHER INFORMATION: EXPRESSED IN HEART, SIGNAL = 48  
OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 81  
OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 28  
OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 24  
OTHER INFORMATION: EXPRESSED IN BT474, SIGNAL = 85  
OTHER INFORMATION: EXPRESSED IN HBL100, SIGNAL = 78  
OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 37  
OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 45  
OTHER INFORMATION: EXPRESSED IN HELA, SIGNAL = 40  
OTHER INFORMATION: EST HUMAN HIT: BE271730.1, EVALUATE 3.00e-50  
OTHER INFORMATION: SWISSPROT HIT: O42161, EVALUATE 4.00e-51  
OTHER INFORMATION: EST\_HUMAN HIT: BE256272.1, EVALUATE 4.00e-50  
US-09-864-761-33414

Query Match 100.0%; Score 75; DB 9; Length 97;  
Best Local Similarity 100.0%; Pred. No. 2.1e-05;  
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 LPHAILRLDLGRDL 15  
| | | | | | | | | | | | | | | | | | | | | |  
Db 50 LPHAILRLDLGRDL 64

RESULT 3  
US-10-424-599-175571  
Sequence 175571, Application US/10424599  
Publication No. US20040031072A1  
GENERAL INFORMATION:  
APPLICANT: La Rosa Thomas J  
APPLICANT: Kovalic David K  
APPLICANT: Zhou Yihua  
APPLICANT: Cao Yongwei  
TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With  
TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement  
FILE REFERENCE: 38-21(53223)B  
CURRENT APPLICATION NUMBER: US/10/424,599  
CURRENT FILING DATE: 2003-04-28  
NUMBER OF SEQ ID NOS: 285684  
SEQ ID NO 175571  
LENGTH: 156  
TYPE: PRT  
ORGANISM: Glycine max  
FEATURE:  
OTHER INFORMATION: Clone ID: PAT\_MRT3847\_129559C.1.pap  
US-10-424-599-175571

Query Match 100.0%; Score 75; DB 15; Length 156;  
Best Local Similarity 100.0%; Pred. No. 3.5e-05;  
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 LPHAILRLDLGRDL 15  
| | | | | | | | | | | | | | | | | | | | | |  
Db 34 LPHAILRLDLGRDL 48

RESULT 4  
US-10-264-049-4044  
Sequence 4044, Application US/10264049  
Publication No. US20040005579A1  
GENERAL INFORMATION:  
APPLICANT: Birse et al.  
TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies  
FILE REFERENCE: PA133p1  
CURRENT APPLICATION NUMBER: US/10/264,049  
CURRENT FILING DATE: 2002-10-04  
PRIOR APPLICATION NUMBER: PCT/US01/18569  
PRIOR FILING DATE: 2001-06-07  
PRIOR APPLICATION NUMBER: US 60/209,467  
PRIOR FILING DATE: 2000-06-07  
NUMBER OF SEQ ID NOS: 4360  
SOFTWARE: PatentIn Ver. 3.1  
SEQ ID NO 4044  
LENGTH: 168  
TYPE: PRT  
ORGANISM: Homo sapiens  
US-10-264-049-4044

Query Match 100.0%; Score 75; DB 15; Length 168;  
Best Local Similarity 100.0%; Pred. No. 3.8e-05;  
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 LPHAILRLDLGRDL 15  
| | | | | | | | | | | | | | | | | | | | | |  
Db 57 LPHAILRLDLGRDL 71

RESULT 5  
US-10-437-963-169247  
Sequence 169247, Application US/10437963  
Publication No. US20040123343A1  
GENERAL INFORMATION:

```
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; APPLICANT: Wu, Wei
; APPLICANT: Boukharov, Andrey A.
; APPLICANT: Barbazuk, Brad
; APPLICANT: Li, Ping
; TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(53221)B
; CURRENT APPLICATION NUMBER: US/10/437,963
; CURRENT FILING DATE: 2003-05-14
; NUMBER OF SEQ ID NOS: 204966
; SEQ ID NO 169247
; LENGTH: 180
; TYPE: PRT
; ORGANISM: Oryza sativa
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT4530_67685C.1.pep
US-10-437-963-169247

Query Match      100.0%; Score 75; DB 16; Length 180;
Best Local Similarity 100.0%; Pred. No. 4.1e-05;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      1 LPHAILRLDLGRDL 15
Db      106 LPHAILRLDLGRDL 120

RESULT 6
US-10-002-631C-56
; Sequence 56, Application US/10002631C
; Publication No. US20030157486A1
; GENERAL INFORMATION:
; APPLICANT: Graff, Jonathon M.
; APPLICANT: Muenster, Matthew
; TITLE OF INVENTION: METHODS TO IDENTIFY SIGNAL SEQUENCES
; FILE REFERENCE: A34943 090495.0243
; CURRENT APPLICATION NUMBER: US/10/002,631C
; CURRENT FILING DATE: 2001-10-31
; PRIOR FILING DATE: 2001-10-31
; PRIOR APPLICATION NUMBER: 60/300,309
; NUMBER OF SEQ ID NOS: 324
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 56
; LENGTH: 219
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-002-631C-56

Query Match      100.0%; Score 75; DB 14; Length 219;
Best Local Similarity 100.0%; Pred. No. 5.1e-05;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      1 LPHAILRLDLGRDL 15
Db      198 LPHAILRLDLGRDL 212

RESULT 7
US-10-767-701-42770
; Sequence 42770, Application US/10767701
; Publication No. US20040172684A1
; GENERAL INFORMATION:
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(53221)B
; CURRENT APPLICATION NUMBER: US/10/767,701
```

```
; CURRENT FILING DATE: 2004-01-29
; NUMBER OF SEQ ID NOS: 63128
; SEQ ID NO 42770
; LENGTH: 230
; TYPE: PRT
; ORGANISM: Sorghum bicolor
; FEATURE:
; OTHER INFORMATION: Clone ID: SORBI-28MAY03-C17410_1.pep
US-10-767-701-42770

Query Match      100.0%; Score 75; DB 16; Length 230;
Best Local Similarity 100.0%; Pred. No. 5.4e-05;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      1 LPHAILRLDLGRDL 15
Db      173 LPHAILRLDLGRDL 187

RESULT 8
US-10-424-599-206389
; Sequence 206389, Application US/10424599
; Publication No. US20040031072A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa Thomas J
; APPLICANT: Kovalic David K
; APPLICANT: Zhou Yihua
; APPLICANT: Cao Yongwei
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(53223)B
; CURRENT APPLICATION NUMBER: US/10/424,599
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 285684
; SEQ ID NO 206389
; LENGTH: 236
; TYPE: PRT
; ORGANISM: Glycine max
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (1)..(236)
; OTHER INFORMATION: unsure at all Xaa locations
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT3847_28396C.1.pep
US-10-424-599-206389

Query Match      100.0%; Score 75; DB 15; Length 236;
Best Local Similarity 100.0%; Pred. No. 5.5e-05;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      1 LPHAILRLDLGRDL 15
Db      41 LPHAILRLDLGRDL 55

RESULT 9
US-10-437-963-162706
; Sequence 162706, Application US/10437963
; Publication No. US20040123343A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; APPLICANT: Wu, Wei
; APPLICANT: Boukharov, Andrey A.
; APPLICANT: Barbazuk, Brad
; APPLICANT: Li, Ping
; TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(53221)B
; CURRENT APPLICATION NUMBER: US/10/437,963
; CURRENT FILING DATE: 2003-05-14
```

```

; NUMBER OF SEQ ID NOS: 204966
; SEQ ID NO 162706
; LENGTH: 238
; TYPE: PRN
; ORGANISM: Oryza sativa
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (1)..(238)
; OTHER INFORMATION: unsure at
; FEATURE:
; OTHER INFORMATION: Clone ID
US-10-437-963-162706

```

```
Query Match      100.0%; Score 75; DB 16; Length 238;
Best Local Similarity 100.0%; Pred. No. 5.6e-05;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

Qy 1 LPFAILRLDLAGRD 15  
Db 24 LPFAILRLDLAGRD 38

```

RESULT 10
US-10-767-701-43575
; Sequence 43575, Application US/10767701
; Publication No. US20040172684A1
; GENERAL INFORMATION:
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof For Plant Improvement
; FILE REFERENCE: 38-21(53535)B
; CURRENT APPLICATION NUMBER: US/10/767,701
; CURRENT FILING DATE: 2004-01-29
; NUMBER OF SEQ ID NOS: 63128
; SEQ ID NO 43575
; LENGTH: 249
; TYPE: PRT
; ORGANISM: Sorghum bicolor
; FEATURE:
; OTHER INFORMATION: Clone ID: SORBI-28MAY03-C5949_1.pep
US-10-767-701-43575

```

```
Query Match      100.0%; Score 75; DB 16; Length 249;
Best Local Similarity 100.0%; Pred. No. 5.9e-05;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

Qy 1 LPHAILRLDLAGRDL 15  
|||  
Db 45 LPHAILRLDLAGRDL 59

```

RESULT 11
US-10-425-114-44494
; Sequence 44494, Application US/10425114
; Publication NO. US20040034888A1
; GENERAL INFORMATION:
; APPLICANT: Liu, Jingdong
; APPLICANT: Zhou, Yihua
; APPLICANT: Kovalic, David K.
; APPLICANT: Screen, Steven E.
; APPLICANT: Tabaska, Jack E.
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-2153313B
; CURRENT APPLICATION NUMBER: US/10/425,114
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 73128
; SEQ ID NO 44494
; LENGTH: 250
; TYPE: PRT

```

```

; ORGANISM: Glycine max
; FEATURE:
; OTHER INFORMATION: Clone ID: 700978095_FLI pep
US-10-425-114-44494

```

```
Query Match      100.0%; Score 75; DB 15; Length 250;
Best Local Similarity 100.0%; Pred. No. 5.9e-05;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

QY 1 LPHAILRLDLGRDL 15  
|||  
Db 46 LPHAILRLDLGRDL 60

RESULT 12  
US-10-425-114-69353  
; Sequence 69353, Application US/10425114  
; Publication No. US20040034888A1  
; GENERAL INFORMATION:  
; APPLICANT: Liu, Jingdong  
; APPLICANT: Zhou, Yihua  
; APPLICANT: Kovalic, David K.  
; APPLICANT: Screen, Steven E  
; APPLICANT: Tabaska, Jack E  
; APPLICANT: Cao, Yongwei  
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With  
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement  
; FILE REFERENCE: 38-21(53313)B  
; CURRENT APPLICATION NUMBER: US/10/425,114  
; CURRENT FILING DATE: 2003-04-28  
; NUMBER OF SEQ ID NOS: 73128  
; SEQ ID NO 69353  
; LENGTH: 254  
; TYPE: PRT  
; ORGANISM: Zea mays  
; FEATURE:  
; OTHER INFORMATION: Clone ID: UC-ZMFLB73022E07\_FLI.pep  
US-10-425-114-69353

```
Query Match          100.0%; Score 75; DB 15; Length 254;
Best Local Similarity 100.0%; Pred. No. 6e-05;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

Qy 1 LP HAILRLDLAGRDL 15  
|||  
50 LP HAILRLDLAGRDL 64  
Db

```

RESULT 13
US-10-424-599-177530
; Sequence 177530, Application US/10424599
; Publication No. US20040031072A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa Thomas J
; APPLICANT: Kovalich David K
; APPLICANT: Zhou Yihua
; APPLICANT: Cao Yongwei
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53223)B
; CURRENT APPLICATION NUMBER: US/10/424,599
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 285684
; SEQ ID NO 177530
; LENGTH: 261
; TYPE: PRT
; ORGANISM: Glycine max
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT3847_131325C.1.pep
US-10-424-599-177530

Query Match 100.0%; Score 75; DB 15; Length 261;
Best Local Similarity 100.0%; Pred. No. 6.2e-05;

```

Query Match 100.0%; Score 75; DB 15; Length 261;  
Best Local Similarity 100.0%; Pred. No. 6.2e-05;

Search completed: April 8, 2005, 12:50:59  
Job time : 42.6667 secs

Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 LPHAILRLDLAAGRL 15  
Db 57 LPHAILRLDLAAGRL 71

RESULT 14  
US-10-425-114-52089  
; Sequence 52089, Application US/10425114  
; Publication No. US2004003488A1  
; GENERAL INFORMATION:  
; APPLICANT: Liu, Jingdong  
; APPLICANT: Zhou, Yihua  
; APPLICANT: Kovalic, David K.  
; APPLICANT: Screen, Steven E  
; APPLICANT: Tabaska, Jack E  
; APPLICANT: Cao, Yongwei  
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With  
; FILE OF INVENTION: Plants and Uses Thereof for Plant Improvement  
; FILE REFERENCE: 38-21(53313)B  
; CURRENT APPLICATION NUMBER: US/10/425,114  
; CURRENT FILING DATE: 2003-04-28  
; NUMBER OF SEQ ID NOS: 73128  
; SEQ ID NO 52089  
; LENGTH: 276  
; TYPE: PRT  
; ORGANISM: Glycine max  
; FEATURE:  
; OTHER INFORMATION: Clone ID: 701149157\_FLI.pep  
US-10-425-114-52089

Query Match 100.0%; Score 75; DB 15; Length 276;  
Best Local Similarity 100.0%; Pred. No. 6.6e-05;  
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 LPHAILRLDLAAGRL 15  
Db 72 LPHAILRLDLAAGRL 86

RESULT 15  
US-10-425-114-57500  
; Sequence 57500, Application US/10425114  
; Publication No. US2004003488A1  
; GENERAL INFORMATION:  
; APPLICANT: Liu, Jingdong  
; APPLICANT: Zhou, Yihua  
; APPLICANT: Kovalic, David K.  
; APPLICANT: Screen, Steven E  
; APPLICANT: Tabaska, Jack E  
; APPLICANT: Cao, Yongwei  
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With  
; FILE OF INVENTION: Plants and Uses Thereof for Plant Improvement  
; FILE REFERENCE: 38-21(53313)B  
; CURRENT APPLICATION NUMBER: US/10/425,114  
; CURRENT FILING DATE: 2003-04-28  
; NUMBER OF SEQ ID NOS: 73128  
; SEQ ID NO 57500  
; LENGTH: 288  
; TYPE: PRT  
; ORGANISM: Zea mays  
; FEATURE:  
; OTHER INFORMATION: Clone ID: UC-ZMROB73004G09\_FLI.pep  
US-10-425-114-57500

Query Match 100.0%; Score 75; DB 15; Length 288;  
Best Local Similarity 100.0%; Pred. No. 6.9e-05;  
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 LPHAILRLDLAAGRL 15  
Db 84 LPHAILRLDLAAGRL 98

**THIS PAGE BLANK (USPTO)**



GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: April 8, 2005, 09:58:56 ; Search time 9.33333 Seconds  
(without alignments)  
154.634 Million cell updates/sec

Title: US-09-423-351C-7  
Perfect score: 75  
Sequence: 1 LPFAILLRLDLGRDL 15

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 283416 seqs, 96216763 residues  
Total number of hits satisfying chosen parameters: 283416

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000  
Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : PIR\_79:\*  
1: pir1:\*  
2: pir2:\*  
3: pir3:\*  
4: pir4:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	75	100.0	92	A31409	actin - California
2	75	100.0	195	S20097	actin 85c - potato
3	75	100.0	325	JC2008	actin homolog prot
4	75	100.0	327	S11452	actin (clone 302)
5	75	100.0	328	S05430	actin beta - grass
6	75	100.0	361	S68089	actin 2 - Arabidop
7	75	100.0	362	A26559	actin type 5, cyto
8	75	100.0	362	S68090	actin 8 - Arabidop
9	75	100.0	370	A29664	actin - sea urchin
10	75	100.0	374	ATBOB	actin beta - bovin
11	75	100.0	374	ATBOG	actin gamma - bovi
12	75	100.0	374	JC5818	gamma-actin - huma
13	75	100.0	375	ATKIC	actin beta - rat
14	75	100.0	375	A49324	actin beta, cytosk
15	75	100.0	375	ATAX	actin - Acanthamo
16	75	100.0	375	ATCHB	actin beta - chick
17	75	100.0	375	ATDO	actin - slime mold
18	75	100.0	375	ATHUB	actin beta - human
19	75	100.0	375	1	actin gamma 1 - hu
20	75	100.0	375	ATMSB	actin beta - mouse
21	75	100.0	375	1	actin gamma - mous
22	75	100.0	375	ATRRB	actin beta, non-mu
23	75	100.0	375	1	actin - maize
24	75	100.0	375	AT1222	actin gamma, cytos
25	75	100.0	375	S71125	actin beta-2, cyto
26	75	100.0	375	S71124	actin beta-1, cyto
27	75	100.0	375	S71126	actin beta, cyto
28	75	100.0	375	A55001	actin beta - goose
29	75	100.0	375	S42103	actin - Puccinia 9

ALIGNMENTS

RESULT 1

A31409 actin - California sea hare (fragments)  
C:Species: Aplysia californica (California sea hare)  
C:Date: 31-Mar-1990 #sequence\_revision 03-Feb-1994 #text\_change 09-Jul-2004  
C:Accession: A31409; A60977  
R:Kennedy, T.E.; Gawinowicz, M.A.; Barzilai, A.; Kandel, E.R.; Sweatt, J.D.  
Proc. Natl. Acad. Sci. U.S.A. 85, 7008-7012, 1988  
A:Title: Sequencing of proteins from two-dimensional gels by using in situ digestion and  
tion in Aplysia.  
A:Reference number: A94207; MUID:88320566; PMID:3413132  
A:Accession: A31409  
A:Molecule type: protein  
A:Residues: 1-18;21-92 <KEN>  
A:Cross-references: UNIPROT:Q7M3Y5  
R:Sweatt, J.D.; Kennedy, T.E.; Wager-Smith, K.; Gawinowicz, M.A.; Barzilai, A.; Karl, K.  
Electrophoresis 10, 152-157, 1989  
A:Title: Development of a database of amino acid sequences for proteins identified and  
A:Reference number: A60977; MUID:89276264; PMID:2731514  
A:Accession: A60977  
A:Molecule type: protein  
A:Residues: 1-20 <SWE>  
C:Superfamily: actin

Query Match 100.0%; Score 75; DB 2; Length 92;  
Best Local Similarity 100.0%; Pred. No. 9.5e-07;  
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 LPFAILLRLDLGRDL 15  
Db 41 LPFAILLRLDLGRDL 55

RESULT 2

S20097 actin 85c - potato (fragment)  
C:Species: Solanum tuberosum (potato)  
C:Date: 22-Nov-1993 #sequence\_revision 17-Oct-1997 #text\_change 09-Jul-2004  
C:Accession: S20097  
R:Drouin, G.; Dover, G.A.  
J. Mol. Evol. 31, 132-150, 1990  
A:Title: Independent gene evolution in the potato actin gene family demonstrated by phy  
A:Reference number: S20092; MUID:91012599; PMID:2120451  
A:Accession: S20097  
A:Molecule type: DNA  
A:Residues: 1-195 <DRO>  
A:Cross-references: UNIPROT:P30170; EMBL:X55747; NID:G21541; PIDN:CAA39277.1; PID:G13455  
C:Genetics:  
A:Introns: 132/1  
C:Superfamily: actin  
C:Keywords: cytoskeleton; structural protein

A;Accession: S05430  
A;Status: translation not shown  
A;Molecule type: DNA  
A;Residues: 1-328 <LIU>  
A;Cross-references: UNIPROT:P83751; EMBL:M25013  
C;Genetics: 41/3; 121/3; 268/1  
C;Superfamily: actin  
C;Keywords: cytoskeleton; methylated amino acid  
F;73/Modified site: 3'-methylhistidine (His) #status predicted

Query Match 100.0%; Score 75; DB 2; Length 328;  
Best Local Similarity 100.0%; Pred. No. 4.1e-06;  
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 LPHAILRLDLGRDL 15  
DB 171 LPHAILRLDLGRDL 185  
|||||

RESULT 6  
S68089  
actin 2 - Arabidopsis thaliana  
C;Species: Arabidopsis thaliana (mouse-ear cress)  
C;Date: 05-Dec-1996 #sequence\_revision 13-Mar-1997 #text\_change 09-Jul-2004  
C;Accession: S68089  
R;McDowell, J.M.; Huang, S.; McKinney, E.C.; An, Y.Q.; Meagher, R.B.  
Genetics 142, 587-602, 1996  
A;Title: Structure and evolution of the actin gene family in Arabidopsis thaliana  
A;Reference number: S68089; MUID:96158109; PMID:8852856  
A;Accession: S68089  
A;Status: not compared with conceptual translation  
A;Molecule type: DNA  
A;Residues: 1-361 <MCD>  
A;Cross-references: UNIPROT:Q93ZL9  
A;Note: mRNA sequencing was also done  
C;Genetics:  
A;Gene: ACT2  
C;Superfamily: actin  
C;Keywords: cytoskeleton; structural protein

Query Match 100.0%; Score 75; DB 2; Length 361;  
Best Local Similarity 100.0%; Pred. No. 4.5e-06;  
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 LPHAILRLDLGRDL 15  
DB 164 LPHAILRLDLGRDL 178  
|||||

RESULT 7  
A26559  
actin type 5, cytosolic - chicken  
C;Species: Gallus gallus (chicken)  
C;Date: 09-Sep-1987 #sequence\_revision 09-Sep-1987 #text\_change 05-Dec-1997  
C;Accession: A26559  
R;Bergsma, D.J.; Chang, K.S.; Schwartz, R.J.  
Mol. Cell. Biol. 5, 1151-1162, 1985  
A;Reference number: A26559; MUID:85213487; PMID:4000121  
A;Accession: A26559  
A;Molecule type: DNA  
A;Residues: 1-362 <BER>  
C;Superfamily: actin  
C;Keywords: cytosol; methylated amino acid  
F;74/Modified site: 3'-methylhistidine (His) #status predicted

Query Match 100.0%; Score 75; DB 2; Length 362;  
Best Local Similarity 100.0%; Pred. No. 4.6e-06;  
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 LPHAILRLDLGRDL 15  
DB 172 LPHAILRLDLGRDL 186  
|||||

```

A;Cross-references: UNIPROT:P60712
A;Note: only peptides that differed in composition from the corresponding peptides of ra
R;Degen, J.L.; Neubauer, M.G.; Degen, S.J.F.; Seyfried, C.E.; Morris, D.R.
J. Biol. Chem. 258, 12153-12162, 1983
A;Title: Regulation of protein synthesis in mitogen-activated bovine lymphocytes. Analys
A;Reference number: A39105; MUID:84032385; PMID:6195151
A;Accession: A39105
A;Molecule type: mRNA
A;Residues: 76-227:344-374 <DEG>
A;Cross-references: GB:K00622; GB:K00623
A;Note: actins beta and gamma were not distinguished in this study
C;Comment: Vertebrate nonmuscle cells contain two highly conserved cytoskeletal actins,
C;Superfamily: actin
C;Keywords: blocked amino end; cell motility; cytoskeleton; methylated amino acid; micro
P;1/Modified site: blocked amino end (Asp) (probably acetylated) #status experimental
P;72/Modified site: 3'-methylhistidine (His) #status predicted

Query Match 100.0%; Score 75; DB 1; Length 374;
Best Local Similarity 100.0%; Pred. No. 4.7e-06;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 LPHAILRLDLAGRDL 15
Db 170 LPHAILRLDLAGRDL 184

RESULT 11
ATBOG
actin gamma - bovine (tentative sequence)
C;Species: Bos primigenius taurus (cattle)
C;Date: 30-Sep-1991 #sequence_revision 30-Sep-1991 #text_change 09-Jul-2004
C;Accession: B14185; A02999
R;Vandekerckhove, J.; Weber, K.
Eur. J. Biochem. 90, 451-462, 1978
A;Title: Actin amino-acid sequences. Comparison of actins from calf thymus, bovine brain
A;Reference number: A14185; MUID:79045349; PMID:213279
A;Accession: B14185
A;Molecule type: protein
A;Residues: 1-374 <VAN>
A;Cross-references: UNIPROT:P02571
A;Note: only peptides that differed in composition from the corresponding peptides of ra
C;Comment: Vertebrate nonmuscle cells contain two highly conserved cytoskeletal actins,
C;Superfamily: actin
C;Keywords: blocked amino end; cell motility; cytoskeleton; methylated amino acid; micro
P;1/Modified site: blocked amino end (Glu) (probably acetylated) #status experimental
P;72/Modified site: 3'-methylhistidine (His) #status predicted

Query Match 100.0%; Score 75; DB 1; Length 374;
Best Local Similarity 100.0%; Pred. No. 4.7e-06;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 LPHAILRLDLAGRDL 15
Db 170 LPHAILRLDLAGRDL 184

RESULT 12
JCS818
gamma-actin - human
C;Species: Homo sapiens (man)
C;Date: 20-Feb-1998 #sequence_revision 13-Mar-1998 #text_change 09-Jul-2004
C;Accession: JC5818; PC4501
R;Hauschildt, S.; Schwarz, C.; Heine, H.; Umer, A.J.; Flad, H.D.; Rietschel, E.T.; Jens
Biochem. Biophys. Res. Commun. 241, 670-674, 1997
A;Title: Actin: A target of lipopolysaccharide-induced phosphorylation in human monocytes
A;Reference number: JC5818; MUID:98096379; PMID:9434766
A;Accession: JC5818
A;Molecule type: protein
A;Residues: 1-374 <HAU>
A;Cross-references: UNIPROT:P02571
A;Experimental source: monocyte
A;Accession: PC4501
A;Molecule type: protein

actin 8 - Arabidopsis thaliana
C;Species: Arabidopsis thaliana (mouse-ear cress)
C;Date: 05-Dec-1996 #sequence_revision 13-Mar-1997 #text_change 09-Jul-2004
C;Accession: S68090
R;McDowell, J.M.; Huang, S.; McKinney, E.C.; An, Y.Q.; Meagher, R.B.
Genetics 142, 587-602, 1996
A;Title: Structure and evolution of the actin gene family in Arabidopsis thaliana.
A;Reference number: S68089; MUID:96158109; PMID:8852856
A;Accession: S68090
A;Status: not compared with conceptual translation
A;Molecule type: DNA
A;Residues: 1-362 <MCD>
A;Cross-references: UNIPROT:O8L894
A;Note: mRNA sequencing was also done
C;Genetics:
A;Gene: ACT8
C;Superfamily: actin
C;Keywords: Cytoskeleton; structural protein

Query Match 100.0%; Score 75; DB 2; Length 362;
Best Local Similarity 100.0%; Pred. No. 4.6e-06;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 LPHAILRLDLAGRDL 15
Db 165 LPHAILRLDLAGRDL 179

RESULT 9
A29664
actin - sea urchin (Strongylocentrotus purpuratus)
C;Species: Strongylocentrotus purpuratus (purple urchin)
C;Date: 15-Dec-1988 #sequence_revision 30-Sep-1991 #text_change 05-Dec-1997
C;Accession: A29664
R;Crain Jr., W.R.; Boshar, M.F.; Cooper, A.D.; Durica, D.S.; Nagy, A.; Steffen, D.
J. Mol. Evol. 25, 37-45, 1987
A;Title: The sequence of a sea urchin muscle actin gene suggests a gene conversion with
A;Reference number: A29664; MUID:87311761; PMID:3114500
A;Accession: A29664
A;Molecule type: DNA
A;Residues: 1-370 <CRA>
A;Cross-references: GB:X05739; GB:X05740; GB:X05741; GB:X05742; GB:X05743
A;Note: the authors translated the codon CAG for residue 260 as Glu
C;Superfamily: actin
C;Keywords: methylated amino acid
P;73/Modified site: 3'-methylhistidine (His) #status predicted

Query Match 100.0%; Score 75; DB 2; Length 370;
Best Local Similarity 100.0%; Pred. No. 4.7e-06;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 LPHAILRLDLAGRDL 15
Db 168 LPHAILRLDLAGRDL 182

RESULT 10
ATBOB
actin beta - bovine (tentative sequence)
C;Species: Bos primigenius taurus (cattle)
C;Date: 30-Sep-1991 #sequence_revision 30-Sep-1991 #text_change 09-Jul-2004
C;Accession: B14185; A39105; A02999; A14185
R;Vandekerckhove, J.; Weber, K.
Eur. J. Biochem. 90, 451-462, 1978
A;Title: Actin amino-acid sequences. Comparison of actins from calf thymus, bovine brain
A;Reference number: A14185; MUID:79045349; PMID:213279
A;Accession: B14185
A;Molecule type: protein
A;Residues: 1-374 <VAN>
```

A;Residues: 1-61;84-112;147-190;196-209;215-253;335-358 <HA2>

A;Experimental source: monocyte

C;Comment: This protein is involved in a signal transduction that eventually leads to m

C;Superfamily: actin

Query Match 100.0%; Score 75; DB 2; Length 374;

Best Local Similarity 100.0%; Pred. No. 4.7e-06;

Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 LPFAILRLDLGRDL 15

Db 170 LPFAILRLDLGRDL 184

RESULT 13

ATATC

actin beta - rat

C;Species: Rattus norvegicus (Norway rat)

C;Date: 13-Jun-1983 #sequence\_revision 13-Jun-1983 #text\_change 22-Jun-1999

C;Accession: A38571; A02999

R;Nudel, U.; Zakut, R.; Shani, M.; Neuman, S.; Levy, Z.; Yaffe, D.

Nucleic Acids Res. 11, 1759-1771, 1983

A;Title: The nucleotide sequence of the rat cytoplasmic beta-actin gene.

A;Reference number: A38571; MUID:83168920; PMID:6300777

A;Accession: A38571

A;Molecule type: DNA

A;Residues: 1-375 <NUD>

A;Cross-references: GB:J00691; NID:G202653; PIDN:AAA40457.1; PID:G202654

C;Comment: Vertebrate nonmuscle cells contain two highly conserved cytoskeletal actins,

C;Genetics:

A;Introns: 41/3; 121/3; 268/1; 328/3

C;Superfamily: actin

C;Keywords: cell motility; cytoskeleton; methylated amino acid; microfilament; mitosis;

F;2-375/Product: actin beta #status predicted <MAT>

F;73/Modified site: 3'-methylhistidine (His) #status predicted

Query Match 100.0%; Score 75; DB 1; Length 375;

Best Local Similarity 100.0%; Pred. No. 4.7e-06;

Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 LPFAILRLDLGRDL 15

Db 171 LPFAILRLDLGRDL 185

RESULT 14

A48324

actin beta, cytoskeletal - common carp

C;Species: Cyprinus carpio (common carp)

C;Date: 03-Feb-1994 #sequence\_revision 11-Apr-1997 #text\_change 09-Jul-2004

C;Accession: A48324

R;Liu, Z.; Zhu, Z.; Roberg, K.; Faras, A.; Guise, K.; Kapuscinski, A.R.; Hackett, P.B.

DNA Seq. 1, 125-136, 1990

A;Title: Isolation and characterization of beta-actin gene of carp (Cyprinus carpio).

A;Reference number: A48324; MUID:92190540; PMID:2134183

A;Accession: A48324

A;Molecule type: DNA

A;Residues: 1-375 <LIU>

A;Cross-references: UNIPROT:P83750; GB:M24113; NID:G213041; PIDN:AAA68886.1; PID:G213042

A;Note: the authors translated the codon TTC for residue 21 as Pro, AAG for residue 50 a

7 as Pro

A;Note: the authors failed to translated the codon GGT for residue 42 as Gly

C;Comment: Vertebrate nonmuscle cells contain two highly conserved cytoskeletal actins,

C;Genetics:

A;Introns: 41/3; 121/3; 268/3; 328/3

C;Superfamily: actin

C;Keywords: acetylated amino end; cell motility; cytoskeleton; methylated amino acid; mi

F;2-375/Product: actin beta, cytoskeletal #status predicted <MAT>

F;72/Modified site: acetylated amino end (Asp) (in mature form) #status predicted

F;73/Modified site: 3'-methylhistidine (His) #status predicted

Query Match

Best Local Similarity 100.0%; Score 75; DB 1; Length 375;

Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 LPFAILRLDLGRDL 15

Db 171 LPFAILRLDLGRDL 185

RESULT 15

ATAX

actin - Acanthamoeba castellanii

C;Species: Acanthamoeba castellanii

C;Date: 31-Dec-1991 #sequence\_revision 31-Dec-1991 #text\_change 09-Jul-2004

C;Accession: A92886; A03004

R;Nellen, W.; Gallwitz, D.

J. Mol. Biol. 159, 1-18, 1982

A;Title: Actin genes and actin messenger RNA in Acanthamoeba castellanii. Nucleotide seq

A;Reference number: A92886; MUID:83033627; PMID:6290670

A;Accession: A92886

A;Molecule type: DNA

A;Residues: 1-375 <NEL>

A;Cross-references: UNIPROT:P02578; GB:V00002; GB:J01016; NID:G5565; PIDN:CAA33399.1; PI

C;Comment: There are at least three actin genes in A. castellanii.

C;Genetics:

A;Introns: 105/3

C;Superfamily: actin

C;Keywords: methylated amino acid

F;73/Modified site: 3'-methylhistidine (His) #status predicted

Query Match 100.0%; Score 75; DB 1; Length 375;

Best Local Similarity 100.0%; Pred. No. 4.7e-06;

Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 LPFAILRLDLGRDL 15

Db 171 LPFAILRLDLGRDL 185

Search completed: April 8, 2005, 10:53:06

Job time : 10.3333 secs

Result No.	Query			ID	Description
	Score	Match	Length		
1	75	100.0	92	2	Q7M3Y5
2	75	100.0	93	2	O6VQP9
3	75	100.0	101	2	O61569
4	75	100.0	104	2	Q1I210
5	75	100.0	109	2	Q28916
6	75	100.0	109	2	Q841A4
7	75	100.0	117	2	Q80012
8	75	100.0	117	2	Q80013
9	75	100.0	125	2	Q802E1
10	75	100.0	125	2	Q802E2
11	75	100.0	128	2	Q852Q7
12	75	100.0	130	2	Q855G0
13	75	100.0	132	2	Q91A84
14	75	100.0	133	2	Q9U4L7
15	75	100.0	133	2	Q9U4L8
16	75	100.0	133	2	Q9U4L9
17	75	100.0	133	2	Q9U4M0
18	75	100.0	136	2	Q95V64
19	75	100.0	140	2	Q68UU0
20	75	100.0	142	2	Q9BGH4
21	75	100.0	145	2	Q9LL42
22	75	100.0	146	2	Q9Y191
23	75	100.0	147	2	O04115
24	75	100.0	150	2	Q68AX1
25	75	100.0	154	2	O48669
26	75	100.0	160	2	Q8T6A7
27	75	100.0	162	2	Q6UBA0
28	75	100.0	165	2	Q8FSF7
29	75	100.0	167	2	Q76D36
30	75	100.0	167	2	Q76D37
31	75	100.0	169	2	Q8SPK6

OS Romalea microptera (Lubber grasshopper).  
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;  
OC Neoptera; Orthopteroidea; Orthoptera; Caelifera; Acridomorpha;  
OC Acridoidea; Romaleidae; Romalea.  
OX NCBI\_TaxID=7007;  
RN [1]  
RP SEQUENCE FROM N.A.  
RA Fei H., Martin T., Jaskowiak K., Hatle J., Borst D.W.;  
RL Submitted (JUN-2003) to the EMBL/GenBank/DBJ databases.  
DR EMBL; AY331667; AAQ24502.1; -.  
FT NON TER 1  
SQ SEQUENCE 93 AA; 10698 MW; 5635F5335A532256 CRC64;  
  
Query Match 100.0%; Score 75; DB 2; Length 93;  
Best Local Similarity 100.0%; Pred. No. 3.8e-06;  
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
  
Qy 1 LPHAILRLDLGRDL 15  
Db 13 LPHAILRLDLGRDL 27  
|||||  
RESULT 3  
O61569 ID O61569 PRELIMINARY; PRT; 101 AA.  
AC O61569;  
DT 01-AUG-1998 (TrEMBLrel. 07, Created)  
DT 01-AUG-1998 (TrEMBLrel. 07, Last sequence update)  
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)  
DE Actin (Fragment).  
GN Name=act;  
OS Osteragia osteragi  
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Strongylida;  
OC Trichostrongyloidea; Haemonchidae; Ostertagiinae; Ostertagia.  
OX NCBI\_TaxID=6317;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=20233682; PubMed=10769181; DOI=10.1042/0264-6021.3470763;  
RA Moore J., Tetley L., Devaney E.;  
RT "Identification of abundant mRNAs from the third stage larvae of the  
parasitic nematode, Osteragia osteragi.";  
RL Biochem. J. 347:763-770(2000).  
DR EMBL; AF052043; AAC06292.1; -.  
DR HSSP; P10983; 1D4X.  
DR GO; GO:0015629; C:actin cytoskeleton; IEA.  
DR GO; GO:0005200; F:structural constituent of cytoskeleton; IEA.  
DR InterPro; IPR004000; Actin\_like.  
DR Pfam; PF00022; Actin; 1.  
FT NON TER 1  
FT NON TER 101  
SQ SEQUENCE 101 AA; 11499 MW; 7AF3A21892AD0D75 CRC64;  
  
Query Match 100.0%; Score 75; DB 2; Length 101;  
Best Local Similarity 100.0%; Pred. No. 4.2e-06;  
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
  
Qy 1 LPHAILRLDLGRDL 15  
Db 8 LPHAILRLDLGRDL 22  
|||||  
RESULT 4  
Q11210 ID Q11210 PRELIMINARY; PRT; 104 AA.  
AC Q11210;  
DT 01-NOV-1996 (TrEMBLrel. 01, Created)  
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)  
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)  
DE Beta-actin (Fragment)  
OS Macaca fascicularis (Crab eating macaque) (Cynomolgus monkey).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Primates; Catarrhini; Cercopitheciidae;  
OC Cercopitheciinae; Macaca.

OX NCBI\_TaxID=9541;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC TISSUE=Lung;  
RA Hino A., Tokuyama Y., Kobayashi M., Yano M., Weir B., Takeda J.,  
RA Bell G.I., MacDonald R.L.;  
RL Submitted (FEB-1995) to the EMBL/GenBank/DBJ databases.  
DR EMBL; U20576; AAA62435.1; -.  
DR HSSP; P10983; 1D4X.  
DR GO; GO:0015629; C:actin cytoskeleton; IEA.  
DR GO; GO:0005200; F:structural constituent of cytoskeleton; IEA.  
DR InterPro; IPR004000; Actin\_like.  
DR Pfam; PF00022; Actin; 1.  
FT NON TER 1  
FT NON TER 104  
SQ SEQUENCE 104 AA; 11438 MW; 3610F3FEFA370C91 CRC64;  
  
Query Match 100.0%; Score 75; DB 2; Length 104;  
Best Local Similarity 100.0%; Pred. No. 4.4e-06;  
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
  
Qy 1 LPHAILRLDLGRDL 15  
Db 42 LPHAILRLDLGRDL 56  
|||||  
RESULT 5  
Q28916 ID Q28916 PRELIMINARY; PRT; 109 AA.  
AC Q28916;  
DT 01-NOV-1996 (TrEMBLrel. 01, Created)  
DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)  
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)  
DE Actin (Fragment).  
OS Macaca fuscata (Japanese macaque).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Primates; Catarrhini; Cercopitheciidae;  
OC Cercopitheciinae; Macaca.  
OX NCBI\_TaxID=9542;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=96071113; PubMed=7577717; DOI=10.1016/0960-0760(95)00157-U;  
RA Yamada-Mouri N., Hirata S., Hayashi M., Kato J.;  
RT "Analysis of the expression and the first exon of aromatase mRNA in  
monkey brain.";  
RL J. Steroid Biochem. Mol. Biol. 55:17-23(1995).  
DR EMBL; S79782; AAB35618.2; -.  
DR HSSP; P10983; 1D4X.  
DR GO; GO:0015629; C:actin cytoskeleton; IEA.  
DR GO; GO:0005200; F:structural constituent of cytoskeleton; IEA.  
DR InterPro; IPR004000; Actin\_like.  
DR Pfam; PF00022; Actin; 1.  
FT NON TER 1  
FT NON TER 109  
SQ SEQUENCE 109 AA; 12333 MW; 64A64E108CDDA54C CRC64;  
  
Query Match 100.0%; Score 75; DB 2; Length 109;  
Best Local Similarity 100.0%; Pred. No. 4.6e-06;  
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
  
Qy 1 LPHAILRLDLGRDL 15  
Db 56 LPHAILRLDLGRDL 70  
|||||  
RESULT 6  
Q841A4 ID Q841A4 PRELIMINARY; PRT; 109 AA.  
AC Q841A4;  
DT 01-JUN-2003 (TrEMBLrel. 24, Created)  
DT 01-JUN-2003 (TrEMBLrel. 24, Last sequence update)  
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)  
DE Actin (Fragment).

```
OS Festuca arundinacea (Tall fescue) (Schedonorus arundinaceus).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Pooidae;
OC Poae; Schedonorus.
RX NCBI_TaxID=4606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Pseudostem;
RA Johnson L.J., Johnson R.D., Scharld C.L., Panaccione D.G.;
RL Submitted (FEB-2003) to the EMBL/GenBank/DBJ databases.
DR EMBL; AY194227; AAC04042.2; -.
DR HSSP; P02577; INMI.
DR GO; GO:0015629; C:actin cytoskeleton; IEA.
DR GO; GO:0005200; F:structural constituent of cytoskeleton; IEA.
DR InterPro; IPR004000; Actin_like.
DR Pfam; PF00022; Actin; 1.
FT NON_TER 1
FT NON_TER 109
FT NON_TER 109
SQ SEQUENCE 109 AA; 12177 MW; 10780BBAS2F6E8EA CRC64;

Query Match 100.0%; Score 75; DB 2; Length 109;
Best Local Similarity 100.0%; Pred. No. 4.6e-06;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 LPHAILRLDLAQRDL 15
Db 52 LPHAILRLDLAQRDL 66

RESULT 7
ID Q80012 PRELIMINARY; PRT; 117 AA.
AC Q80012;
DT 01-JUN-2003 (TrEMBLrel. 24, Created)
DT 01-JUN-2003 (TrEMBLrel. 24, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Beta-actin 2 (Fragment).
GN Name=ACTB2;
OS Hyla japonica (Japanese tree frog).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Amphibia; Batrachia; Anura; Neobatrachia; Hyloidea; Hylidae; Hylinae;
OC Hyla.
OX NCBI_TaxID=109175;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Liver;
RA Kohno S., Kamishima Y., Iguchi T.;
RL Submitted (SEP-2002) to the EMBL/GenBank/DBJ databases.
CC -!- FUNCTION: Actins are highly conserved proteins that are involved
CC in various types of cell motility and are ubiquitously expressed
CC in all eukaryotic cells (By similarity).
CC -!- SIMILARITY: Belongs to the actin family.
CC EMBL; AB092519; BAC66207.1; -.
DR HSSP; P60712; 1HLU.
DR GO; GO:0015629; C:actin cytoskeleton; IEA.
DR GO; GO:0005200; F:structural constituent of cytoskeleton; IEA.
DR InterPro; IPR004000; Actin; 1.
DR Pfam; PF00022; Actin; 1.
DR PRINTS; PR00190; ACTIN.
DR PROSITE; PS01132; ACTINS_ACT_LIKE; 1.
KW Structural protein.
FT NON_TER 1
FT NON_TER 117
FT NON_TER 117
SQ SEQUENCE 117 AA; 13291 MW; 64C6BAC4F84184CF CRC64;

Query Match 100.0%; Score 75; DB 2; Length 117;
Best Local Similarity 100.0%; Pred. No. 5e-06;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 LPHAILRLDLAQRDL 15
Db 92 LPHAILRLDLAQRDL 106

RESULT 9
ID Q802E1 PRELIMINARY; PRT; 125 AA.
AC Q802E1;
DT 01-JUN-2003 (TrEMBLrel. 24, Created)
DT 01-JUN-2003 (TrEMBLrel. 24, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Beta actin (Fragment).
GN Name=ACT-B;
OS Zoarces viviparus (Belout).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;
OC Acanthomorpha; Acanthopterygii; Perciformes; Zoarcoidei;
OC Zoarcidae; Zoarces.
OX NCBI_TaxID=48416;
RN [1]
RP SEQUENCE FROM N.A.
RA Lucassen M., Sokolov E., Poertner H.-O.;
RL Submitted (JAN-2003) to the EMBL/GenBank/DBJ databases.
CC -!- FUNCTION: Actins are highly conserved proteins that are involved
CC in various types of cell motility and are ubiquitously expressed
CC in all eukaryotic cells (By similarity).
CC -!- SIMILARITY: Belongs to the actin family.
CC EMBL; AY227658; AAC07272.1; -.
DR HSSP; P60712; 1HLU.
DR GO; GO:0015629; C:actin cytoskeleton; IEA.
DR GO; GO:0005200; F:structural constituent of cytoskeleton; IEA.
```

```
RESULT 8
Q80013 PRELIMINARY; PRT; 117 AA.
ID Q80013;
AC Q80013;
DT 01-JUN-2003 (TrEMBLrel. 24, Created)
DT 01-JUN-2003 (TrEMBLrel. 24, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Beta-actin 1 (Fragment).
GN Name=ACTB1;
OS Hyla japonica (Japanese tree frog).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Amphibia; Batrachia; Anura; Neobatrachia; Hyloidea; Hylidae; Hylinae;
OC Hyla.
OX NCBI_TaxID=109175;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Liver;
RA Kohno S., Kamishima Y., Iguchi T.;
RL Submitted (SEP-2002) to the EMBL/GenBank/DBJ databases.
CC -!- FUNCTION: Actins are highly conserved proteins that are involved
CC in various types of cell motility and are ubiquitously expressed
CC in all eukaryotic cells (By similarity).
CC -!- SIMILARITY: Belongs to the actin family.
CC EMBL; AB092519; BAC66207.1; -.
DR HSSP; P60712; 1HLU.
DR GO; GO:0015629; C:actin cytoskeleton; IEA.
DR GO; GO:0005200; F:structural constituent of cytoskeleton; IEA.
DR InterPro; IPR004000; Actin; 1.
DR Pfam; PF00022; Actin; 1.
DR PRINTS; PR00190; ACTIN.
DR PROSITE; PS01132; ACTINS_ACT_LIKE; 1.
KW Structural protein.
FT NON_TER 1
FT NON_TER 117
FT NON_TER 117
SQ SEQUENCE 117 AA; 13291 MW; 64C6BAC4F84184CF CRC64;

Query Match 100.0%; Score 75; DB 2; Length 117;
Best Local Similarity 100.0%; Pred. No. 5e-06;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 LPHAILRLDLAQRDL 15
Db 92 LPHAILRLDLAQRDL 106

RESULT 9
ID Q802E1 PRELIMINARY; PRT; 125 AA.
AC Q802E1;
DT 01-JUN-2003 (TrEMBLrel. 24, Created)
DT 01-JUN-2003 (TrEMBLrel. 24, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Beta actin (Fragment).
GN Name=ACT-B;
OS Zoarces viviparus (Belout).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;
OC Acanthomorpha; Acanthopterygii; Perciformes; Zoarcoidei;
OC Zoarcidae; Zoarces.
OX NCBI_TaxID=48416;
RN [1]
RP SEQUENCE FROM N.A.
RA Lucassen M., Sokolov E., Poertner H.-O.;
RL Submitted (JAN-2003) to the EMBL/GenBank/DBJ databases.
CC -!- FUNCTION: Actins are highly conserved proteins that are involved
CC in various types of cell motility and are ubiquitously expressed
CC in all eukaryotic cells (By similarity).
CC -!- SIMILARITY: Belongs to the actin family.
CC EMBL; AY227658; AAC07272.1; -.
DR HSSP; P60712; 1HLU.
DR GO; GO:0015629; C:actin cytoskeleton; IEA.
DR GO; GO:0005200; F:structural constituent of cytoskeleton; IEA.
```

```

DR InterPro; IPR004000; Actin_like.
DR Pfam; PF00022; Actin; 1.
DR PRINTS; PR00190; ACTIN.
DR PROSITE; PS01132; ACTINS_ACT_LIKE; 1.
DR Structural protein.
FT NON_TER 1
FT NON_TER 125
SQ SEQUENCE 125 AA; 14100 MW; 7B33EA4CF5EA8B0 CRC64;

Query Match 100.0%; Score 75; DB 2; Length 125;
Best Local Similarity 100.0%; Pred. No. 5.4e-06;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 LPHAILRLDLAGRDL 15
| | | | | | | | | | | | | | |
Db 85 LPHAILRLDLAGRDL 99

RESULT 10
Q802E2 PRELIMINARY; PRT; 125 AA.
ID Q802E2 AC Q802E2;
DT 01-JUN-2003 (TrEMBLrel. 24, Created)
DT 01-JUN-2003 (TrEMBLrel. 24, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Beta actin (Fragment).
GN Name=ACT-B;
OS Pachycara brachycephalum (Antarctic eelpout) (Austrolycichthys brachycephalus).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;
OC Acanthomorpha; Acanthopterygii; Percomorpha; Perciformes; Zeacridoidei;
OC Zoarcidae; Pachycara.
OX NCBI_TaxID=36221;
RN [1]
RP SEQUENCE FROM N.A.
RA Lucassen M., Sokolov E., Foertner H.-O.;
RL Submitted (JAN-2003) to the EMBL/GenBank/DBJ databases.
CC -1- FUNCTION: Actins are highly conserved proteins that are involved in various types of cell motility and are ubiquitously expressed in all eukaryotic cells (By similarity).
CC -1- SIMILARITY: Belongs to the actin family.
DR HSP; AY227657; AA072721.1; -.
DR HGSP; P60712; 1HLU.
DR GO; GO:0015629; C:actin cytoskeleton; IEA.
DR GO; GO:0005200; F:structural constituent of cytoskeleton; IEA.
DR InterPro; IPR004000; Actin_like.
DR Pfam; PF00022; Actin; 1.
DR PRINTS; PR00190; ACTIN.
DR PROSITE; PS01132; ACTINS_ACT_LIKE; 1.
DR Structural protein.
FT NON_TER 1
FT NON_TER 125
SQ SEQUENCE 125 AA; 14100 MW; 7B33EA4CF5EA8B0 CRC64;

Query Match 100.0%; Score 75; DB 2; Length 125;
Best Local Similarity 100.0%; Pred. No. 5.4e-06;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 LPHAILRLDLAGRDL 15
| | | | | | | | | | | | | | |
Db 85 LPHAILRLDLAGRDL 99

RESULT 11
Q852Q7 PRELIMINARY; PRT; 128 AA.
ID Q852Q7 AC Q852Q7;
DT 01-JUN-2003 (TrEMBLrel. 24, Created)
DT 01-JUN-2003 (TrEMBLrel. 24, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Beta actin (Fragment).
GN Name=MpACT1;

```



```

GO: GO:0005200; F:structural constituent of cytoskeleton; IEA.
DR InterPro; IPR004000; Actin_like.
DR Pfam; PF00022; Actin; 1.
DR PRINTS; PRO0190; ACTIN.
DR PROSITE; PS01132; ACTINS_ACT_LIKE; 1.
KW Structural protein.
FT NON_TER 1
FT NON_TER 133
SQ SEQUENCE 133 AA; 14955 MW; 27DFC1FF30D4984F CRC64;

Query Match 100.0%; Score 75; DB 2; Length 133;
Best Local Similarity 100.0%; Pred. No. 5.8e-06;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 LPFALLRLDLGRDL 15
DB 83 LPFALLRLDLGRDL 97
|||||
|||||

RESULT 15
QYU4L8 PRELIMINARY; PRT; 133 AA.
ID AC Q9U4L8;
AC Q9U4L8;
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE Actin 1 (Fragment).
OS Agniaspis citricola.
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC Neoptera; Endopterygota; Hymenoptera; Apocrita; Chalcidoidea;
OC Encyrtidae; Encyrtinae; Agniaspis.
OC NCBI_TaxID=105762;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Taiwan population;
RA Hoy M.A., Jeyaprasath A., Morakote R., Lo P.K.C., Nguyen R.;
RT "Genomic analyses of two populations of Agniaspis citricola
(RHymenoptera: Encyrtidae) suggest that a cryptic species may exist.";
RT Biol. Control 17:1-10(2000).
CC -!- FUNCTION: Actins are highly conserved proteins that are involved
CC in various types of cell motility and are ubiquitously expressed
CC in all eukaryotic cells (By similarity).
CC -!- SIMILARITY: Belongs to the actin family.
CC EMBL; AF164630; AAF22278.1; -.
DR HSSP; P10983; 1D4X.
DR GO; GO:0015629; C:actin cytoskeleton; IEA.
DR GO; GO:0005200; F:structural constituent of cytoskeleton; IEA.
DR InterPro; IPR004000; Actin_like.
DR Pfam; PF00022; Actin; 1.
DR PRINTS; PRO0190; ACTINS.
DR PROSITE; PS01132; ACTINS_ACT_LIKE; 1.
KW Structural protein.
FT NON_TER 1
FT NON_TER 133
SQ SEQUENCE 133 AA; 14941 MW; 62CB94EF37A39F3F CRC64;

Query Match 100.0%; Score 75; DB 2; Length 133;
Best Local Similarity 100.0%; Pred. No. 5.8e-06;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 LPFALLRLDLGRDL 15
DB 83 LPFALLRLDLGRDL 97
|||||
|||||

Search completed: April 8, 2005, 12:03:10
Job time : 58.8667 secs

```

**THIS PAGE BLANK (USPTO)**